

Documentation for all Data Sets

HEALTH ABC DATA ANALYSIS FILE.....	2
PARTICIPANT HISTORY FILE (PH).....	3
YEAR 4 CLINIC VISIT DATA (Y4ClnVis).....	5
YEAR 4 CORE HOME VISIT DATA (Y4CoreHV).....	7
YEAR 4 PROXY CONTACT DATA (Y4PROXY).....	8
42-MONTH SEMI-ANNUAL TELEPHONE CONTACT DATA (SA42Mo).....	10
42-MONTH PROXY CONTACT DATA (SA42PROX).....	11
MISSED FOLLOW-UP CONTACT DATA (MissVis).....	12
SPECIAL MISSING VALUE CODES.....	13
DROPPED VARIABLES.....	15
LISTINGS.....	15
DATA DICTIONARY (datadict.xls, datadict.dat).....	16
Appendix I DROPPED VARIABLES AND ALTERNATES.....	17
Appendix II BUA READING CENTER VARIABLE LIST.....	27
Appendix III 42-MONTH CALCULATED VARIABLES.....	29

HEALTH ABC DATA ANALYSIS FILE

To use the data, please contact the PI at your site.

Contents:

SAS Datasets

Y4ClnVis	Year 4 Clinic Visit and Return Visit data
Y4CoreHV	Year 4 Core Home Visit Workbook, Year 3 Return Visit Phlebotomy and Year 4 Return Visit Blood Gas and Laboratory Processing data related to home visits, visit specific workbook?
Y4Proxy	Year Proxy Interview data and associated Year 4 Proxy Home Visit Workbook, Year 4 Clinic Visit Workbook, and/or Year 4 Return Visit Phlebotomy and Laboratory Processing data related to Proxy Visits
SA42Mo	42-month Semi-Annual Telephone Contact Data
SA42Prox	42-month Proxy Interview data
MissVis	Missed Follow-up Contact data
Y4Read	Year 4 Reading Center data
Y4Calc	Year 4 calculated (derived) variables
BMDNotes	Explanatory notes for BMD QC code variables

In addition the following files, not specific to any year but updated each time data are released, can be found at the top of the Current Datasets listing on the Health ABC website under the Current Datasets link:

PH	Participant History File
Formats	SAS Format Library

Substudies (see Substudy Documentation)

Y4WtChg	Year 4 weight change substudy
Y45FLU	Year 4&5 flu substudy

DXA datasets for Repeated Measure Analysis

Dxrean14	Whole body BMD data for baseline (Year 1), Year 2, Year 3, and Year 4 for <u>year-by-year longitudinal analyses only</u> . Where Year 2, 3, or 4 scans required a change of region of interest (ROI), preceding scans have been reanalyzed to match the Year 2, 3, or 4 ROI (see Reading Center documentation).
----------	---

PARTICIPANT HISTORY FILE (PH)

1. General description

The PH file contains general information about the participants enrolled in the study. Variables included are:

HABCID	Health ABC Enrollment ID# without the 2-letter prefix
HCFAID	HCFA Screening ID (as assigned by the Coordinating Center)
DOB	Date of Birth
DOD	Date of Death
GENDER	Gender (1=Male; 2=Female)
RACE	Race (1=White; 2=Black)
SITE	Clinic Site (1=Memphis, 2=Pittsburgh)
CV1AGE	Age at Year 1 Clinic Visit
CV1DATE	Year 1 Clinic Visit Date
SV06AGE	Age at 6-Month Contact
SV06DATE	6-Month Contact Date
VITAL06M	Vital Status at time of 6-month contact
CV2AGE	Age at Year 2 Clinic Visit
CV2DATE	Year 2 Clinic Visit Date
VITAL12M	Vital Status at time of Year 2 (12-month) contact
SV18AGE	Age at 18-Month Contact
SV18DATE	18-Month Contact Date
VITAL18M	Vital Status at time of 18-month contact
CV3AGE	Age at Year 3 Clinic Visit
CV3DATE	Year 3 Clinic Visit Date
VITAL24M	Vital Status at time of Year 3 (24-month) contact
SV30AGE	Age at 30-Month Contact
SV30DATE	30-Month Contact Date
VITAL30M	Vital Status at time of 30-month contact
CV4AGE	Age at Year 4 Clinic Visit
CV4DATE	Year 4 Clinic Visit Date
VITAL36M	Vital Status at time of Year 4 (36-month) contact
SV42AGE	Age at 42-Month Contact
SV42DATE	42-Month Contact Date
VITAL42M	Vital Status at time of 42-month contact
VERSIONPH	Participant History File Release Date
VStatus	Vital Status (1=Alive, 2=Dead) as of PH file release date

The birthdate, race, and gender data come from the edited HCFA data. The Year 1 clinic visit date was taken from the corrected final participant list provided by the each clinic (with later corrections as necessary). All other dates were taken from the corresponding Clinic Visit Workbook, Core Home Visit workbook, Proxy Interview or Proxy Contact Home Visit Workbook, or Semi-Annual Telephone Contact form; participants who missed a visit have no corresponding date (set to .A). Age at each contact is a calculated variable based on birthdate and that contact date.

There are 3075 observations in the PH file. The demographic breakdown of participants in this dataset is as follows:

African-American Female	729
African-American Male	552
White Female	855
White Male	939
Memphis Participants	1548
Pittsburgh Participants	1527

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under PH) or by searching the Datadict file (sort by form or database).

3. Dataset structure and contents

The PH file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
HCFAID	HCFA ID (as assigned by the Coordinating Center)

4. Condition of data

a. Known data errors: None at this time. The data have been edited. Editing will, however, be ongoing (iterative), so use of the most recent dataset is always advised.

b. Strength and weaknesses of dataset items: If a death has been reported on a Missed Visit Form, an Event Form, or the Report of Death, the participant is listed as deceased in the vital status variable (VStatus) variable. Note: VStatus is vital status **as of the release date of the participant history file**, not as of any particular visit. This variable should only be used to determine the most up-to-date snapshot of vital status in Health ABC as of the file date. Vital status variables (VITALxxM) have been created for each Health ABC contact as follows: If a participant had a particular contact, or if they missed a contact but they were determined to still be alive at the time their contact was due (participant refused the contact, was too ill, etc), then VITALxxM is alive. If they missed a contact and were later discovered to have died before the end of their contact window, then VITALxxM is dead. If they missed a contact and were later discovered to have died after the end of their contact window, the VITALxxM is alive for that contact, but dead for the next. Finally, if they missed a contact without a determination of their vital status (ppt could not be located, withdrew, etc) and no further contact with vital status determination has been made since then, then VITALxxM is missing.

The date of death (DOD) variable represents the best available information about the date of death for deceased participants as of the creation date of participant history file (PH).^{*} If a Report of Death form has been entered for the participant, the date of death from that adjudication form is used. If there is no Report of Death form yet, this information is taken from the Event Form dataset, and is therefore an un-confirmed, un-adjudicated date of death.

* Run proc contents in SAS to see creation date of the PH file.

c. **Missing Value Conventions:** See Special Missing Value Codes on page 14 for special missing value codes applied

5. Dataset index formulation and key variable mapping

The PH file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

YEAR 4 CLINIC VISIT DATA (Y4ClnVis)

1. General description

The Y4ClnVis file contains information about the participants enrolled in the study gathered from the Year 4 Clinic Visit Workbook, Year 4 Questionnaire, and the Return Visit forms associated with clinic visits. If a participant did not have a Year 4 clinic visit, a Year 4 home visit, or a Year 4 proxy visit, they should have a Missed Follow-up Contact form that explains why. In addition, the variable VISITYPE (which replaces the old variables COREHV and MISSVIS) has been appended to allow the analyst to account for all participants, whether or not they had a Year 4 clinic visit. If VISITYPE=1 or 2, the participant's Year 4 data can be found in Y4CoreHV. If VISITYPE=6 or 8, the participant's Year 4 data can be found in Y4PROXY. If VISITYPE=3, 4 or 6, the participant did not have a Year 4 contact and their Missed Follow-up Contact data can be found in MISSVIS.

There are 3075 observations in the Y4ClnVis file.

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under Y4ClnVis) or by searching the Datadict file (sort by form or database). Variable names can also be found on the annotated forms. Please note that not all variables on the forms are contained in the dataset. All variables not found in the dataset are listed in Dropped Variables and Alternates (Appendix I). Alternate variables to use (if applicable) are also listed.

3. Dataset structure and contents

The Y4ClnVis file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh*
VISITYPE	(0=Clinic, 1=Home, 2=Phone, 3=Missed (other), 4=Deceased, 5=Withdrew, 6=Proxy Home, 8=Proxy Phone; 10=Mixed Visit)
MISSREAS	Reason Y4 Visit Missed

* Must link to Health ABC participant history file (PH) to add this variable.

4. Condition of data

a. Known data errors:

Year 4 Questionnaire

There has been some inconsistency in the responses to Question 24 (regular care or assistance to a child or disabled or sick adult). For those who provide care on a 24/7 basis, some participant have included time that the care receiver was asleep, giving an answer of $24 \times 7 = 168$, while others estimate the number of hours they are actually caring for the person. This was addressed in part by a Q&A in Year 6 (Q&A Ref #264), in which examiners were directed to only count time that the participant is actively, physically caring for the child. Analysts may wish to reduce these extreme values by an estimate of the number of sleeping hours. This problem has existed with this question from baseline.

Year 4 Clinic Visit Workbook

Examiners were somewhat inconsistent in their recording of the ankle-arm blood pressure measurements. In most cases when “unable to locate tibial artery” was marked, the corresponding tibial systolic measurements were missing. However, in some cases, this value was completed as 0. Similarly, there were several cases where the tibial systolic measurements were marked 0, but only a few of them also had “unable to locate tibial artery” marked. The tibial systolic measurements should only be 0 if the tibial artery pulse was located before cuff inflation and then never heard again after complete deflation of the cuff. No attempt has been made in this version of the data to deal with this inconsistency as instructions from the prime mover were not received in time. Eventually, the calculated variables for ankle-arm blood pressure (see calculated variable documentation) will use decision rules for each of these cases.

b. Strength and weaknesses of dataset items:

Two new variables, D320TM1S and D320TM2S have been added, which convert the 20-meter walk times to seconds. When both this variable and the corresponding steps variable (D320STP1 or D320STP2, respectively) are 0, the participant tried, but was unable to complete the 20-m walk.

Calculated (derived) variables have now been created to complement those created for Years 1, 2, and 3 data. To avoid confusion, these variables are listed in Y4Calc.

A new value (10) has been added to the variable VISITYPE to account for strange cases when a participant had a “mixed” type of visit. In Year 4 this happened with one participant (HB5097), who had a home visit, then came into the clinic and did all of the measurements except the walks and phlebotomy (which were done during the home visit). Thus, this participant’s data will be found in Y4ClnVis for everything except walks, phlebotomy and processing, and other measurements not done at the Year 4 clinic visit (radial pulse, knee crepitus, isometric strength, ultrasound). Walk, phlebotomy and processing data will be found in Y4COREHV. Data for measurements done at both the home visit and the clinic visit were dropped from Y4COREHV for this participant to avoid confusion.

Knee X-ray Data

During Year 4, only new knee pain cases (or old, existing knee pain who had never had an x-ray) were eligible for knee x-ray. The data from the Knee X-ray reading center are not included in this dataset. Rather, they will be included in the next release of the Knee OA substudy data.

c. Missing Value Conventions: See Special Missing Value Codes on page 14 for special missing value codes applied.

5. Dataset index formulation and key variable mapping

The Y4ClnVis file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

YEAR 4 CORE HOME VISIT DATA (Y4CoreHV)

1. General description

The Y4CoreHV file contains information about the participants enrolled in the study gathered from the revised Core Home Visit Workbook, plus Return Visit Phlebotomy and Laboratory Processing forms associated with a Year 4 home visit. If a participant did not have a Year 4 clinic visit, a Year 4 proxy contact, or a Year 4 home visit, they should have a Missed Follow-up Contact form that explains why. These data can be found the dataset MissVis (see page 12).

There are 308 observations in the Y4CoreHV file.

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under Y4CoreHV) or by searching the Datadict file (sort by form or database). Variable names can also be found on the annotated forms. Please note that not all variables on the forms are contained in the dataset. All variables not found in the dataset are listed in Dropped Variables and Alternates (Appendix I). Alternate variables to use (if applicable) are also listed.

3. Dataset structure and contents

The Y4CoreHV file contains a single observation per participant and has records only for those participants who had a “home visit.” A participant may have “home visit” data either from a true visit in the home (ZCCONTAC=1), or from a telephone or other contact in lieu of a home visit (ZCCONTAC=2 or 3), if the participant was unwilling or unable to have an examiner come to the home. If the data were collected by telephone, all physical measurements (data collected on pages 34-48; i.e., all variables beginning with Z4) will be missing. In addition, if participants were unwilling to answer the entire questionnaire on pages 2-28, examiners were instructed to concentrate on the questions marked with a star. Thus, some participants may be missing un-starred variables, and the exact variables completed or missing may vary from participant to participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh [*]

^{*} Must link to Health ABC participant history file (PH) to add this variable.

4. Condition of data

a. Known data errors: No known errors at this time.

Strength and weaknesses of dataset items: The BUA Reading Center dataset, which applies only to this subset of participants, has been appended (see Appendix II for variable list).

A variable Z4WTK (=2.2*Z4WTLBS) has been added for easy comparison to D2WTK in Y4ClnVis.

Calculated (derived) variables have been created to complement those created for Years 1, 2 and 3 data. To avoid confusion, these variables are listed in Y4Calc.

Data for HB5097 are included in Y4CoreHV although they are also represented in Y4ClnVis. This is because they had a home visit, then came into the clinic and did all of the measurements except the walks and phlebotomy (which were done during the home visit). Thus, this participant's data will be found in Y4ClnVis for everything except walks, phlebotomy and processing, and other measurements not done at the Year 4 clinic visit (radial pulse, knee crepitus, isometric strength, ultrasound). Walk, phlebotomy and processing data will be found in Y4COREHV. Data for measurements done at both the home visit and the clinic visit were dropped from Y4COREHV for this participant to avoid confusion.

c. Missing Value Conventions: See Special Missing Value Codes page 14 for special missing value codes applied.

5. Dataset index formulation and key variable mapping

The Y4CoreHV file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

YEAR 4 PROXY CONTACT DATA (Y4PROXY)

1. General description

As the Health ABC cohort has aged, some participants have begun having cognitive difficulties that prevent their being able to answer interview questions, and in some cases, being able to understand the instructions for some measurements. In response to this situation, two new forms were created near the end of Year 3, the Proxy Interview and the Proxy Contact Home Visit Workbook. A participant may have one of three possible types of proxy contacts:

- a proxy phone visit (49 participants in Year 4 fall into this group), in which case they will have only Proxy Interview data;
- a proxy home visit (14 participants), in which case they will have both a Proxy Interview and a Proxy Contact Home Visit Workbook; or
- a proxy clinic visit (6 participants), in which case they would have a Proxy Interview and a Year 4 Clinic Visit Workbook.

Data for this last group will be found in two places: interview data will be found in Y4Proxy, while physical measurements will be found in Y4ClnVis. Note that a “proxy phone” contact is empirically

defined as having interview data only, but no physical measurements, regardless of where the participant or their proxy was at the time of the interview. This definition was used for the creation of the VISITYPE variable, but was not always followed by the clinics in assigning a value to YACONTAC. Thus VISITYPE is a more reliable variable to use to determine how many proxy contact have only interview information, vs how many have physical measurements as well.

There are 69 observations in the Y4Proxy file.

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under Y4Proxy) or by searching the Datadict file (sort by form or database). Calculated variable names and descriptions are listed in the documentation for the Calculated Variables. Variable names can also be found on the annotated forms. Please note that not all variables on the forms are contained in the dataset. All variables not found in the dataset are listed in Dropped Variables and Alternates (Appendix I). Alternate variables to use (if applicable) are also listed.

3. Dataset structure and contents

The Y4Proxy file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh [*]

4. Condition of data

a. Known data errors: None at this time.

b. Strength and weaknesses of dataset items: No calculated variables have been created yet for this dataset. The only standard calculated variables that might apply to these data are the self-reported function variables; however, it has been reported that proxies over-report functional limitation (Elam, et al. Am J Public Health. 1991; 81:1127), and therefore more consideration needs to be applied to how these variables should be calculated.

Because the proxy home visit weight may be measured in pounds or kilograms, the variables YCWT and YCLBSKG have been replaced by a new variable YCWTK (weight in kg) for convenience.

c. Missing Value Conventions: See Special Missing Value Codes page 14 for special missing value codes applied

5. Dataset index formulation and key variable mapping

The Y4Proxy file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

^{*} Must link to Health ABC participant history file (PH) to add this variable.

42-MONTH SEMI-ANNUAL TELEPHONE CONTACT DATA (SA42Mo)

1. General description

Data from the 42-month visit are now complete and fully edited, and therefore are included in this data release. The SA42Mo file contains information about the participants enrolled in the study gathered from the 42-month Semi-Annual Telephone Contact forms. In some cases, a participant was unable to complete the contact by telephone because of deafness, cognitive impairment, or other reasons, and the information was gathered from a proxy. Data for these participants can be found in SA42PROX (see page 11). If a participant did not have a 42-month follow-up telephone call or a proxy contact, they should have a Missed Follow-up Contact form that explains why. These data can be found the dataset MissVis (see page 12). All participants are accounted for in this dataset (n=3075) by including VISITYPE (see below).

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under SA42Mo) or by searching the Datadict file (sort by form or database). Calculated variable names and descriptions are listed in Appendix III. Variable names can also be found on the annotated forms. Please note that not all variables on the forms are contained in the dataset. All variables not found in the dataset are listed in Dropped Variables and Alternates (Appendix I). Alternate variables to use (if applicable) are also listed.

3. Dataset structure and contents

The SA42Mo file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh*
VISITYPE	(3=Missed (other), 4=Deceased, 5=Withdrew, 7=SATC Phone, 8=Proxy Phone;)
MISSREAS	Reason 42-month contact missed

4. Condition of data

a. Known data errors: None at this time.

b. Strength and weaknesses of dataset items: Calculated variables have been added to the dataset (see Appendix III).

c. Missing Value Conventions: See Special Missing Value Codes page 14 for special missing value codes applied

5. Dataset index formulation and key variable mapping

The SA42Mo file is sorted by HABCID, which is a unique identifier for each participant.

* Must link to Health ABC participant history file (PH) to add this variable.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

42-MONTH PROXY CONTACT DATA (SA42PROX)

1. General description

As discussed above (page 8) some participants experience either transient or permanent cognitive impairment to a degree that precludes their being able to answer questions reliably, and a proxy is needed. In addition, for the semi-annual phone contacts, sometimes a participant is hard of hearing and is unable to answer the questions, despite being totally unimpaired cognitively. Thus, 45 participants have a Proxy Interview in lieu of a 42-month telephone interview.

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under SA42Prox) or by searching the Datadict file (sort by form or database). Variable names can also be found on the annotated forms. Please note that not all variables on the forms are contained in the dataset. All variables not found in the dataset are listed in Dropped Variables and Alternates (Appendix I). Alternate variables to use (if applicable) are also listed.

3. Dataset structure and contents

The SA42Prox file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh [*]

4. Condition of data

a. Known data errors: No known errors at this time

b. Strength and weaknesses of dataset items: No calculated variables have been created yet for this dataset. The only standard calculated variables that might apply to these data are the self-reported function variables; however, it has been reported that proxies over-report functional limitation (Elam, et al. Am J Public Health. 1991; 81:1127), and therefore more consideration needs to be applied to how these variables should be calculated.

c. Missing Value Conventions: See Special Missing Value Codes page 14 for special missing value codes applied

5. Dataset index formulation and key variable mapping

The SA42Prox file is sorted by HABCID, which is a unique identifier for each participant.

^{*} Must link to Health ABC participant history file (PH) to add this variable.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

MISSED FOLLOW-UP CONTACT DATA (MissVis)

1. General description

The MissVis file contains information about the participants who have missed a follow-up contact (died, refused, lost to follow-up, etc) at any time during the study, up through the 42-month contact. If a participant missed an annual or semi-annual contact, they should have a Missed Follow-up Contact form that explains why. The number of Missed Follow-up Contact forms related to each of these visits is shown below

6-month follow-up telephone call	23
Year 2 clinic/home visit	77
18-month semi-annual telephone call	137
Year 3 clinic/home/proxy visit	94
30-month semi-annual telephone call	198
Year 4 clinic/home/proxy visit	160
42-month semi-annual telephone call	201

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under MissVis) or by searching the Datadict file (sort by form or database). Variable names can also be found on the annotated forms.

3. Dataset structure and contents

The MissVis file contains multiple observations per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh*
BJID2	Contact missed (9=6-month, 2=Year 2, 10=18-month, 3=Year 3, 11=30-month, 4=Year 4, 12=42-month)

4. Condition of data

a. Known data errors: None.

b. Strength and weaknesses of dataset items: If a participant missed a visit due to death or withdrawal from the study, the Missed Follow-up Contact corresponding to the first contact missed for this reason is usually the last Missed Follow-up Contact for that participant. That is, field centers were instructed not to continue completing Missed Follow-up Contacts for each subsequent contact missed after the death of a participant or their withdrawal from the study. If a participant could not

* Must link to Health ABC participant history file (PH) to add this variable.

be located at one contact and therefore had a Missed Follow-up Contact completed for that contact, then subsequently was found to have died before that contact, the death was recorded on a Missed Follow-up Contact form for the subsequent contact. That is, the Missed Follow-up Contact information reflects the status of the participant as known to the field center at the time of the scheduled contact. Missed Follow-up Contact data should not be used to determine approximate date of death, nor even numbers of participants who had died as of a particular follow-up contact. The best information available at the time of the data analysis file (10/2/02) regarding date of death can be found in the Participant History file (PH, DOD, see page 4). If a participant was found to have both a Missed Follow-up Contact form for a particular contact and the corresponding contact forms (e.g. Clinic Visit Workbook, Core Home Visit Workbook, Proxy Interview, Proxy Contact Home Visit Workbook, or Semi-Annual Follow-Up Contact form), the Missed Follow-up Contact form data for that participant were deleted from the analysis file.

c. Missing Value Conventions: See Special Missing Value Codes page 14 for special missing value codes applied

5. Dataset index formulation and key variable mapping

The MissVis file is sorted by HABCID, which is a unique identifier for each participant. The combination of HABCID and BJID2 is a unique identifier for a participant/contact record in this dataset.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets. The MissVis file must first be subsetted by BJID2 to the contact desired before merging with a contact-specific, one-record-per-participant dataset.

SPECIAL MISSING VALUE CODES

SAS allows for stratification of missing values. The following missing values have been assigned:

. = 'Missing Form'
.A = 'A:Not Applicable'
.E = 'E:Recoded to Missing'
.F = 'F:Variable Missing from Form'
.L = 'L:Permanently Lost'
.M = 'M:Missing'
.N = 'N:Not Required'
.T = 'T:Missing Due to Technical Problems' (reading center data only)
.U = 'Unacceptable'

Description

.: Missing Form

Used when a value is missing because the entire form has not been entered.

A: Not Applicable

Used when a value is missing but the value is not required (due to simple skip pattern logic)

E: Recoded to Missing

Used to flag that a value was entered originally but should not have been (due to a skip pattern logic error) and that the value has been recoded to missing

F: Variable Missing from Form

Used to flag a variable that was not originally on the form (form was revised during the visit year) and therefore there is no value for this participant

L: Permanently Lost

Used to flag a tracking variable (e.g., D2WB, D3RL3) when a measurement involving a Reading Center was done, but the data either never made it to the Reading Center, or was lost at the Reading Center. A list of lost measurements, along with a brief description of what happened, can be found in Appendix I of the Reading Center Dataset documentation). Every effort was made to track down these data before they were declared "permanently lost" and the flag assigned.

M:Missing

Used to flag missing values when the value is required (i.e., true missing values).

N:Not Required

Used when a value is missing but the value is not required (not due to simple skip pattern logic). For example, for checkbox variables which are "Check all that apply" each one, individually, is not required. In these cases, a summary calculated variable (not included on the dataset) was used to edit missing responses. Some variables whose skip pattern logic is non-standard (i.e., the skip pattern involves several variables and forms) also have .N flags when missing, whether or not a response was required due to the skip pattern.

T:Missing Due to Technical Problems

Used when a value is missing from the Reading Center dataset due to technical difficulties. An explanation of when this value has been assigned can be found under Strengths and Weaknesses of (Reading Center) dataset items for each Reading Center (Reading Center data documentation; Substudy documentation).

U:Unacceptable

Used with certain Reading Center data when the data exist but cannot be used. These include:

- DXA data when the whole scan has been reviewed as unacceptable
- PMNC data when the QC process at the Eastern Virginia Medical School determined that part or all of a PMNC file could not be used for various reasons
- ECG change data when a suppression code invalidates the comparison between years of some or all aspects of a tracing (See Reading Center Documentation)

General Strategies for Using Special Missing Values

In SAS, when using special missing values in logical expressions, the missing value is no longer only equal to '.' To express a value equal to missing, the code should be written: <= .z or alternately: le .z

To express a value not equal to missing, the code should be written >.z or alternately: gt .z
.Z is the greatest value of missing available in SAS.

DROPPED VARIABLES

A number of variables appear on the annotated forms but will not be found in the datasets. These variables are listed in Appendix I. There are several reasons why variables were dropped:

1. Participant confidentiality: identifying information such as participant name, address, telephone number, etc. are omitted from the analysis file. All participants are instead identified by the HABC Enrollment ID# (HABCID).
2. Bookkeeping variables: a number of variables were put on the forms merely for bookkeeping purposes and are extremely unlikely to be useful for analysis. If an investigator notices that such a variable has been dropped and believes it should not have been, s/he should feel free to contact the Coordinating Center (HABCHelp@psg.ucsf.edu) and let us know that it should be included in future datasets.
3. Redundant variables: in many cases the same information was collected numerous times. We have made a special effort to clean up one version of each of these; and to avoid analysis using uncleaned data, the uncleaned version is omitted.

Appendix I lists the omitted variables in alphabetical order. There are two columns; the first, entitled "Variable omitted," lists the name of the variable not included in the dataset. The second column, entitled "Variable to use," lists the cleaned variable, for redundant variables, or the reason the variable was not included.

LISTINGS

A PDF listing of the SAS proc contents printout for all SAS datasets can be found under the "Proc Contents for All Datasets" link.

A text file, Formats.lst, showing all formats and value descriptions (e.g.: 1=White, 2=Black) contained in the SAS Format Library can also be downloaded. Click on FormatsList under the List of Current Datasets on the Health ABC website.

The following files are zipped together in a self-extracting document also available on the Health ABC website under the Current Datasets listing

Datadict.xls Microsoft Excel 4.0 spreadsheet containing information about all variables included in the SAS datasets

Datadict.dat Tab-delimited text file containing the same information as above.

DATA DICTIONARY (datadict.xls, datadict.dat)

General description

This is a searchable/sortable file that contains all the variables included in the SAS datasets. The following fields are included:

Variable	Variable Name
Label	Description of the variable
Form	Form or reading center origin of the variable
Page	Page number (not applicable for reading center or PH data)
Database	Database location of the variable
Variable Types	Type of variable (text, categorical (numeric), continuous (numeric), date, time, etc.)
Possible Values	Range of possible values associated with the variable (used for range edits)
SAS Format	SAS format assigned to the variable
Page Order	Variable order on the data collection forms (useful for sorting)

General Strategies for Use

The data dictionary is provided in two formats: Excel 4.0 and tab-delimited text. They contain exactly the same information. The tab-delimited file was generated to provide easy access to those who do not have Excel version 4.0 or higher.

The file is currently sorted by page order, form, page. It can be used to search for details on a particular variable, to group by database or form, or to find a variable location on a form. It is a little more user friendly than the standard SAS proc contents, as it can be sorted, easily searched, and it provides additional details such as page number and form.

Appendix I

DROPPED VARIABLES AND ALTERNATES

Variable omitted	Variable to use
BJCONTAC	BJID2
BJVISIT	BJID2
BLACROS	N/A (confidentiality)
BLID	HABCID
BLLINK	N/A (bookkeeping)
BLREF13A	N/A (bookkeeping)
BLREF13B	N/A (bookkeeping)
BLREF13C	N/A (bookkeeping)
BLREF14A	N/A (bookkeeping)
BLREF14B	N/A (bookkeeping)
BLREF14C	N/A (bookkeeping)
BLREF15A	N/A (bookkeeping)
BLREF15B	N/A (bookkeeping)
BLREF15C	N/A (bookkeeping)
BLREF16A	N/A (bookkeeping)
BLREF16B	N/A (bookkeeping)
BLREF16C	N/A (bookkeeping)
BLREF17A	N/A (bookkeeping)
BLREF17B	N/A (bookkeeping)
BLREF17C	N/A (bookkeeping)
BLREF18A	N/A (bookkeeping)
BLREF18B	N/A (bookkeeping)
BLREF18C	N/A (bookkeeping)
BLREF19A	N/A (bookkeeping)
BLREF19B	N/A (bookkeeping)
BLREF19C	N/A (bookkeeping)
BLREF19D	N/A (bookkeeping)
BLREF19E	N/A (bookkeeping)
BLREF19F	N/A (bookkeeping)
BLREF20A	N/A (bookkeeping)
D1ACROS	N/A (confidentiality)
D1DOC	N/A (bookkeeping)
D1ID	HABCID
D1LINK	N/A (bookkeeping)
D1TIME1	N/A (bookkeeping)
D1TIME2	N/A (bookkeeping)
D2ACROS	N/A (confidentiality)
D2CVAFH	FHCV (Y4Read)
D2CVAPF	PFCV (Y4Read)
D2DSAMP	AAMP (Y4Read)

Variable omitted	Variable to use
D2DSAMP1	N/A (bookkeeping)
D2FHAMP	FHAMP (Y4Read)
D2FHAMP1	N/A (bookkeeping)
D2ID	N/A (bookkeeping)
D2LINK	N/A (bookkeeping)
D2PFAMP	PFAMP (Y4Read)
D2PFAMP1	N/A (bookkeeping)
D2SCAN	N/A (bookkeeping)
D2SCDTE	N/A (bookkeeping)
D2VPAV	VIBAVG (Y4Read)
D2VPBV	VIBADAP (Y4Read)
D2VPS10	N/A (bookkeeping)
D2VPSTP	N/A (bookkeeping)
D2VPVAR	VIBVAR (Y4Read)
D320TM1A	D320TIM1 (min) or D320TM1S (sec)
D320TM1B	D320TIM1 (min) or D320TM1S (sec)
D320TM2A	D320TIM2 (min) or D320TM2S (sec)
D320TM2B	D320TIM2 (min) or D320TM2S (sec)
D34TIMEA	D34TIME
D34TIMEB	D34TIME
D3ACROS	N/A (confidentiality)
D3AVTORQ	KCTMEAN (Y4Read)
D3DIAB2	D2DIA2
D3ID	HABCID
D3LINK	N/A (bookkeeping)
D3PKTORQ	KCTMAX (Y4Read)
D3SYB2	D2SY2
D4ACROS	N/A (confidentiality)
D4AMPM4	D4VTM24
D4AMPM5	D4BDTM24
D4AMPMS	D4TMSP24
D4BLDRM	D4BDTM24
D4ID	HABCID
D4LINK	N/A (bookkeeping)
D4LMAPM	D4MHM24
D4MHM	D4MHM24
D4TIMESP	D4TMSP24
D4VTM	D4VTM24
D5ACROS	N/A (confidentiality)
D5AMPM4	D5VTM24
D5AMPM5	D5BDTM24
D5BLDRM	D5BDTM24
D5ID	HABCID
D5LINK	N/A (bookkeeping)
D5LMAPM	D5MHM24

Variable omitted	Variable to use
D5MHM	D5MHM24
D5VTM	D5VTM24
D6ACROS	N/A (confidentiality)
D6AMPMS	D6TMS24
D6ID	HABCID
D6TIMESP	D6TMS24
DAACROS	N/A (confidentiality)
DAEWHRS	DAEWTIM
DAEWMINS	DAEWTIM
DAHCHRS	DAHCTIM
DAHCMINS	DAHCTIM
DAHIA1HR	DAH1TIME
DAHIA1MN	DAH1TIME
DAHIA2HR	DAH2TIME
DAHIA2MN	DAH2TIME
DAID	HABCID
DALINK	N/A (bookkeeping)
DAREF34A	N/A (bookkeeping)
DAREF34B	N/A (bookkeeping)
DAREF34C	N/A (bookkeeping)
DAREF35A	N/A (bookkeeping)
DAREF35B	N/A (bookkeeping)
DAREF35C	N/A (bookkeeping)
DAREF36A	N/A (bookkeeping)
DAREF36B	N/A (bookkeeping)
DAREF36C	N/A (bookkeeping)
DAREF37A	N/A (bookkeeping)
DAREF37B	N/A (bookkeeping)
DAREF37C	N/A (bookkeeping)
DAREF38A	N/A (bookkeeping)
DAREF38B	N/A (bookkeeping)
DAREF38C	N/A (bookkeeping)
DAREF39A	N/A (bookkeeping)
DAREF39B	N/A (bookkeeping)
DAREF39C	N/A (bookkeeping)
DAREF40A	N/A (bookkeeping)
DAREF40B	N/A (bookkeeping)
DAREF40C	N/A (bookkeeping)
DAREF40D	N/A (bookkeeping)
DAREF40E	N/A (bookkeeping)
DAREF40F	N/A (bookkeeping)
DAREF41A	N/A (bookkeeping)
DBACROS	N/A (confidentiality)
DBADDRES	N/A (confidentiality)
DBADDYN	N/A (bookkeeping)

Variable omitted	Variable to use
DBAPT	N/A (confidentiality)
DBC1APT	N/A (confidentiality)
DBC1CITY	N/A (confidentiality)
DBC1FNAM	N/A (confidentiality)
DBC1LNAM	N/A (confidentiality)
DBC1NOK	N/A (bookkeeping)
DBC1PHON	N/A (confidentiality)
DBC1POA	N/A (bookkeeping)
DBC1REL	N/A (bookkeeping)
DBC1STAT	N/A (confidentiality)
DBC1STRT	N/A (confidentiality)
DBC1YN	N/A (bookkeeping)
DBC1ZIP	N/A (confidentiality)
DBC2APT	N/A (confidentiality)
DBC2CITY	N/A (confidentiality)
DBC2FNAM	N/A (confidentiality)
DBC2LNAM	N/A (confidentiality)
DBC2NOK	N/A (bookkeeping)
DBC2PHON	N/A (confidentiality)
DBC2POA	N/A (bookkeeping)
DBC2REL	N/A (bookkeeping)
DBC2STAT	N/A (confidentiality)
DBC2STRT	N/A (confidentiality)
DBC2ZIP	N/A (confidentiality)
DBCIAPT	N/A (confidentiality)
DBCICITY	N/A (confidentiality)
DBCIFNAM	N/A (confidentiality)
DBCILNAM	N/A (confidentiality)
DBCINOK	N/A (bookkeeping)
DBCIPOA	N/A (bookkeeping)
DBCIREL	N/A (bookkeeping)
DBCISTAT	N/A (confidentiality)
DBCISTRT	N/A (confidentiality)
DBCITELE	N/A (confidentiality)
DBCITY	N/A (confidentiality)
DBCIYN	N/A (confidentiality)
DBCIZIP	N/A (confidentiality)
DBDCITY	N/A (confidentiality)
DBDFNAME	N/A (confidentiality)
DBDLNAME	N/A (confidentiality)
DBDPHONE	N/A (confidentiality)
DBDSTATE	N/A (confidentiality)
DBDSTRT	N/A (confidentiality)
DBDZIP	N/A (confidentiality)
DBFNAME	N/A (confidentiality)

Variable omitted	Variable to use
DBID	HABCID
DBK2APT	N/A (confidentiality)
DBK2CITY	N/A (confidentiality)
DBK2FNAM	N/A (confidentiality)
DBK2LNAM	N/A (confidentiality)
DBK2PHON	N/A (confidentiality)
DBK2REL	N/A (bookkeeping)
DBK2STAT	N/A (confidentiality)
DBK2STRT	N/A (confidentiality)
DBK2ZIP	N/A (confidentiality)
DBKAPT	N/A (confidentiality)
DBKCITY	N/A (confidentiality)
DBKFNAM	N/A (confidentiality)
DBKLNAM	N/A (confidentiality)
DBKPHONE	N/A (confidentiality)
DBKREL	N/A (bookkeeping)
DBKSTATE	N/A (confidentiality)
DBKSTRT	N/A (confidentiality)
DBKYN	N/A (bookkeeping)
DBKZIP	N/A (confidentiality)
DBLINK	N/A (bookkeeping)
DBLNAME	N/A (confidentiality)
DBMAAPT	N/A (confidentiality)
DBMACITY	N/A (confidentiality)
DBMADATE	N/A (confidentiality)
DBMASTAT	N/A (bookkeeping)
DBMASTRT	N/A (confidentiality)
DBMATELE	N/A (confidentiality)
DBMAZIP	N/A (confidentiality)
DBP2APT	N/A (confidentiality)
DBP2CITY	N/A (confidentiality)
DBP2FNAM	N/A (confidentiality)
DBP2LNAM	N/A (confidentiality)
DBP2PHON	N/A (confidentiality)
DBP2REL	N/A (confidentiality)
DBP2STAT	N/A (confidentiality)
DBP2STRT	N/A (confidentiality)
DBP2YN	N/A (bookkeeping)
DBP2ZIP	N/A (confidentiality)
DBPAAPT	N/A (confidentiality)
DBPACITY	N/A (confidentiality)
DBPAFNAM	N/A (confidentiality)
DBPALNAM	N/A (confidentiality)
DBPAPHON	N/A (confidentiality)
DBPAREL	N/A (bookkeeping)

Variable omitted	Variable to use
DBPASTAT	N/A (confidentiality)
DBPASTRT	N/A (confidentiality)
DBPAYN	N/A (bookkeeping)
DBPAZIP	N/A (confidentiality)
DBPHONE	N/A (confidentiality)
DBSTATE	N/A (confidentiality)
DBSTREET	N/A (confidentiality)
DBWKPHONE	N/A (confidentiality)
DBZIP	N/A (confidentiality)
YAACROS	N/A (confidentiality)
YAADDRESS	N/A (confidentiality)
YAAPT	N/A (confidentiality)
YACITY	N/A (confidentiality)
YACONTAC	N/A (confidentiality)
YADCITY	N/A (confidentiality)
YADFNNAME	N/A (confidentiality)
YADLNAME	N/A (confidentiality)
YADPHONE	N/A (confidentiality)
YADSTATE	N/A (confidentiality)
YADSTRT	N/A (confidentiality)
YADZIP	N/A (confidentiality)
YAFNAME	N/A (confidentiality)
Y Aid	HABCID
YALINK	N/A (bookkeeping)
YALNAME	N/A (confidentiality)
YAMAAPT	N/A (confidentiality)
YAMACITY	N/A (confidentiality)
YAMADATE	N/A (confidentiality)
YAMASTATE	N/A (confidentiality)
YAMASTRT	N/A (confidentiality)
YAMATELE	N/A (confidentiality)
YAMAZIP	N/A (confidentiality)
YAMOVE	N/A (confidentiality)
YAMOVE2	N/A (confidentiality)
YAPHONE	N/A (confidentiality)
YAREF11A	N/A (bookkeeping)
YAREF11B	N/A (bookkeeping)
YAREF11C	N/A (bookkeeping)
YAREF12A	N/A (bookkeeping)
YAREF12B	N/A (bookkeeping)
YAREF12C	N/A (bookkeeping)
YAREF13A	N/A (bookkeeping)
YAREF13B	N/A (bookkeeping)
YAREF13C	N/A (bookkeeping)
YAREF14A	N/A (bookkeeping)

Variable omitted	Variable to use
YAREF14B	N/A (bookkeeping)
YAREF14C	N/A (bookkeeping)
YAREF15A	N/A (bookkeeping)
YAREF15B	N/A (bookkeeping)
YAREF15C	N/A (bookkeeping)
YAREF16A	N/A (bookkeeping)
YAREF16B	N/A (bookkeeping)
YAREF16C	N/A (bookkeeping)
YAREF17A	N/A (bookkeeping)
YAREF17B	N/A (bookkeeping)
YAREF17C	N/A (bookkeeping)
YAREF17D	N/A (bookkeeping)
YAREF17E	N/A (bookkeeping)
YAREF17F	N/A (bookkeeping)
YAREF18A	N/A (bookkeeping)
YASTATE	N/A (bookkeeping)
YASTREET	N/A (bookkeeping)
YAVISIT	N/A (bookkeeping)
YAWKPHON	N/A (confidentiality)
YAZIP	N/A (confidentiality)
YBACROS	N/A (confidentiality)
YBVISIT	N/A (bookkeeping)
YCACROS	N/A (confidentiality)
YCID	HABCID
YCLBSKG	YCWTK
YCLINK	N/A (bookkeeping)
YCVISIT	N/A (bookkeeping)
YCWT	YCWTK
Z4BUQUI1	QUI
Z4BUBUA1	BUA
Z4BUSOS1	SOS
Z4BUAST1	N/A (bookkeeping)
Z4BUQUI2	QUI
Z4BUBUA2	BUA
Z4BUSOS2	SOS
Z4BUAST2	N/A (bookkeeping)
Z4BUQUI3	QUI
Z4BUBUA3	BUA
Z4BUAST3	N/A (bookkeeping)
Z4BUDIF1	N/A (bookkeeping)
Z4BUDIF2	N/A (bookkeeping)
Z4BU2AST	N/A (bookkeeping)
Z4ID	HABCID
Z4ACROS	N/A (confidentiality)
Z4LINK	N/A (bookkeeping)

Variable omitted	Variable to use
Z4TYPE	ZBTYPE
Z4SCAN1	WBSCANID, WBMODE
Z4SCAN2	HIPSCNID
Z4SCDTE1	WBDATE
Z4SCDTE2	HIPDATE
ZBID	HABCID
ZBACROS	N/A (confidentiality)
ZBTIME1	N/A (bookkeeping)
ZBTIME2	N/A (bookkeeping)
ZBTYPE	N/A (bookkeeping)
ZCID	HABCID
ZCACROS	N/A (confidentiality)
ZCLINK	N/A (bookkeeping)
ZCTYPE	ZBTYPE
ZCREF23A	N/A (bookkeeping)
ZCREF23B	N/A (bookkeeping)
ZCREF23C	N/A (bookkeeping)
ZCREF24A	N/A (bookkeeping)
ZCREF24B	N/A (bookkeeping)
ZCREF24C	N/A (bookkeeping)
ZCREF25A	N/A (bookkeeping)
ZCREF25B	N/A (bookkeeping)
ZCREF25C	N/A (bookkeeping)
ZCREF26A	N/A (bookkeeping)
ZCREF26B	N/A (bookkeeping)
ZCREF26C	N/A (bookkeeping)
ZCREF27A	N/A (bookkeeping)
ZCREF27B	N/A (bookkeeping)
ZCREF27C	N/A (bookkeeping)
ZCREF28A	N/A (bookkeeping)
ZCREF28B	N/A (bookkeeping)
ZCREF28C	N/A (bookkeeping)
ZCREF29A	N/A (bookkeeping)
ZCREF29B	N/A (bookkeeping)
ZCREF29C	N/A (bookkeeping)
ZCREF29D	N/A (bookkeeping)
ZCREF29E	N/A (bookkeeping)
ZCREF29F	N/A (bookkeeping)
ZCREF30A	N/A (bookkeeping)
ZCEWHRS	ZCEWTIM
ZCEWMINS	ZCEWTIM
ZDTYPE	ZBTYPE
ZDLINK	N/A (bookkeeping)
ZDDFNAME	N/A (confidentiality)
ZDDLNAME	N/A (confidentiality)

Variable omitted	Variable to use
ZDDSTRT	N/A (confidentiality)
ZDDCITY	N/A (confidentiality)
ZDDZIP	N/A (confidentiality)
ZDDSTATE	N/A (confidentiality)
ZDDPHONE	N/A (confidentiality)
ZDFNAME	N/A (confidentiality)
ZDLNAME	N/A (confidentiality)
ZDSTREET	N/A (confidentiality)
ZDAPT	N/A (confidentiality)
ZDCITY	N/A (confidentiality)
ZDZIP	N/A (confidentiality)
ZDSTATE	N/A (confidentiality)
ZDPHONE	N/A (confidentiality)
ZDWKPHON	N/A (confidentiality)
ZDMASTRT	N/A (confidentiality)
ZDMAAPT	N/A (confidentiality)
ZDMACITY	N/A (confidentiality)
ZDMAZIP	N/A (confidentiality)
ZDMASTAT	N/A (confidentiality)
ZDADDRES	N/A (confidentiality)
ZDMATELE	N/A (confidentiality)
ZDMADATE	N/A (bookkeeping)
ZDCIFNAM	N/A (confidentiality)
ZDCILNAM	N/A (confidentiality)
ZDCISTR	N/A (confidentiality)
ZDCIAPT	N/A (confidentiality)
ZDCICITY	N/A (confidentiality)
ZDCI STAT	N/A (confidentiality)
ZDCI ZIP	N/A (confidentiality)
ZDCI TELE	N/A (confidentiality)
ZDCI REL	N/A (confidentiality)
ZDCI NOK	N/A (bookkeeping)
ZDCI POA	N/A (bookkeeping)
ZDC1FNAM	N/A (confidentiality)
ZDC1LNAM	N/A (confidentiality)
ZDC1STR	N/A (confidentiality)
ZDC1APT	N/A (confidentiality)
ZDC1CITY	N/A (confidentiality)
ZDC1STAT	N/A (confidentiality)
ZDC1ZIP	N/A (confidentiality)
ZDC1PHON	N/A (confidentiality)
ZDC1REL	N/A (confidentiality)
ZDC1NOK	N/A (bookkeeping)
ZDC1POA	N/A (bookkeeping)
ZDC2FNAM	N/A (confidentiality)

Variable omitted	Variable to use
ZDC2LNAM	N/A (confidentiality)
ZDC2STRT	N/A (confidentiality)
ZDC2APT	N/A (confidentiality)
ZDC2CITY	N/A (confidentiality)
ZDC2STAT	N/A (confidentiality)
ZDC2ZIP	N/A (confidentiality)
ZDC2PHON	N/A (confidentiality)
ZDC2REL	N/A (confidentiality)
ZDC2NOK	N/A (bookkeeping)
ZDC2POA	N/A (bookkeeping)
ZDKFNAME	N/A (confidentiality)
ZDKLNAME	N/A (confidentiality)
ZDKSTRT	N/A (confidentiality)
ZDKAPT	N/A (confidentiality)
ZDKCITY	N/A (confidentiality)
ZDKSTATE	N/A (confidentiality)
ZDKZIP	N/A (confidentiality)
ZDKPHONE	N/A (confidentiality)
ZDKREL	N/A (confidentiality)
ZDPAFNAM	N/A (confidentiality)
ZDPALNAM	N/A (confidentiality)
ZDPASTRT	N/A (confidentiality)
ZDPAAPT	N/A (confidentiality)
ZDPACITY	N/A (confidentiality)
ZDPASTAT	N/A (confidentiality)
ZDPAZIP	N/A (confidentiality)
ZDPAPHON	N/A (confidentiality)
ZDPAREL	N/A (confidentiality)

Appendix II

BUA READING CENTER VARIABLE LIST

Variable Name	Variable Description	Variable Label	Value Label
BUA	BUA mean	BUA MEAN	
BUACV	CV for BUA tests	CV FOR BUA TESTS	
BUADATE	Exam date	EXAM DATE	MMDDYY
BUAMAX	Maximum BUA value	MAXIMUM BUA VALUE	
BUAMIN	Minimum BUA value	MINIMUM BUA VALUE	
BUARNGE	Range of BUA values	RANGE OF BUA VALUES	
BUASIDE	BUA side	BUA SIDE	L=left R=right
BUASOURC	Source of data	SOURCE OF DATA	1=Sahara 2=Form
BUASTD	St. Dev. Of BUA tests	ST. DEV. OF BUA TESTS	
EBMD	Estimated BMD mean	ESTIMATED BMD MEAN	
EBMDCV	CV for estimated BMD	CV FOR EST BMD	
EBMDMAX	Max value for estimated BMD	MAX VALUE FOR EST BMD	
EBMDMIN	Min value for estimated BMD	MIN VALUE FOR EST BMD	
EBMDRNGE	Range of estimated BMD values	RANGE OF EST BMD VALUES	
EBMDSTD	Std dev for estimated BMD	STD DEV FOR EST BMD	
FLAGBUA	Flag set to 1 when ppt has BUA data	Flag set to 1 when ppt has BUA data	1 = ppt has data
NBUA	# Of BUA tests	# OF BUA TESTS	
NEBMD	# Of est BMD tests	# OF EST BMD TESTS	
NQUI	# Of QUI tests	# OF QUI TESTS	
NSOS	# Of speed of sound tests	# OF SOS TESTS	
QUI	Stiffness (QUI) mean	STIFFNESS (QUI) MEAN	
QUICV	CV for QUI	CV FOR QUI	
QUIMAX	Max value for QUI	MAX VALUE FOR QUI	
QUIMIN	Min value for QUI	MIN VALUE FOR QUI	
QUIRNGE	Range of QUI values	RANGE OF QUI VALUES	
QUISTD	Std dev for QUI	STD DEV FOR QUI	
SOS	Speed of sound mean	SPEED OF SOUND MEAN	m/s
SOSCV	CV for speed of sound	CV FOR SOS	m/s

Variable Name	Variable Description	Variable Label	Value Label
SOSMAX	Max value for speed of sound	MAX VALUE FOR SOS	m/s
SOSMIN	Min value for speed of sound	MIN VALUE FOR SOS	m/s
SOSRNGE	Range of speed of sound values	RANGE OF SOS VALUES	m/s
SOSSTD	Std dev for speed of sound	STD DEV FOR SOS	m/s

Appendix III

42-MONTH CALCULATED VARIABLES

Variable	Variable Description	Grouping	Page
CSAINDEX	Climbing stairs ability index	Self-Reported Function Calculated Vars.	31
EASE1F	Ease climbing 1 flight	Self-Reported Function Calculated Vars.	31
EASE1M	Ease walking 1 mile	Self-Reported Function Calculated Vars.	31
EASE2F	Ease climbing 2 flights	Self-Reported Function Calculated Vars.	31
EASEQM	Ease walking 1/4 mile	Self-Reported Function Calculated Vars.	30
WKAINDEX	Walking ability index	Self-Reported Function Calculated Vars.	31

42-Month Self-Reported Function Calculated Variables

Investigator Name: Eleanor Simonsick

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASEQM	Ease walking 1/4 mile	Recode of BLDWQMEZ where 3 is easiest; imputations applied for missing BLDWQMEZ	If BLDWQMYN in (0,7,8) then EASEQM=7-BLDWQMEZ If BLDWQMYN=1 then EASEQM=4-BLDWQMDF	If BLDWQMYN=9 and (BLMNRS>0 or BLMNRS4 in (1,2)) then BLDWQMYN=1; If BLDWQMEZ=8 and BLDWQMYN>8 then BLDWAQMDF=2; If BLDWQMEZ=8 and BLDWQMYN>8 then BLDWAQMEZ=2; If BLDWQMYN in (0,7,8) and BLDWQMEZ<0 then EASEQM=4; If (BLDWQMYN=0 and BLDWQMEZ<0) then EASEQM=7-BLDW1MEZ; if (BLDWQMYN=0 and BLDWQMEZ<0) and BLDW1MEZ<0 and (BLDW1MYN=1 or BLDW1MYN=8 or BLDW1MYN†.Z) then EASEQM=4; if (BLDWQMYN=0 and BLDWQMEZ<0) and (BLDW1MYN=0 and BLDW1MEZ†.Z) then EASEQM=5 if BLDWQMYN=1 and BLDWQNDF<0 then EASEQM=7-BLDWQMEZ; if BLDWQMYN=1 and BLDWQNDF<0 and BLDWQMEZ<0 then EASEQM=2; if BLDWQMYN<0 and BLDWQMDF>0 then EASEQM=4-BLDWQMDF; if BLDWQMYN<0 and BLDWQMEZ>0 then EASEQM=7-BLDWQMEZ; if BLDWQMYN=8 and BLD2QMEZ=8 then EASEQM=4; if BLDWQMYN=9 and BLMNRS<0 and (BLMNRS4<0 or BLMNRS4=8) then EASEQM=.	6=very easy 5=somewhat easy 4=not that easy 3=a little difficult 2=somewhat difficult 1=very difficult 0=unable to do

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE1M	Ease walking 1 mile	Recode of BLDW1MEZ where 3 is easiest; imputations applied for missing BLDW1MEZ	If BLDW1MYN=1 then EASE1M=0; if BLDW1MEZ=3 then EASE1M=1; if BLDW1MEZ=2 then EASE1M=2; if BLDW1MEZ=1 then EASE1M=3	If BLDW1MYN=8 and (BLDW1MEZ<0 or BLDW1MEZ=8) then EASE1M=1; if BLDW1MYN†.Z and BLDW1MEZ†.Z then EASE1M=EASEQM-4; if BLDW1MYN=0 and BLDW1MEZ†.Z then EASE1M=EASEQM-4; if BLDWQMYN=1 then EASE1M=0; if EASEQM > 0 and (EASEQM-3 < EASE1M) and EASEQM†3 then EASE1M=0, else If EASEQM-3<EASE1M then EASE1M=EASEQM-3; If BLDW1MYN=8 and BLDW1MEZ in (1,2,3) then EASE1M=4-BLDW1MEZ; If BLDW1MYN=0 and BLDW1MEZ=8 then EASE1M=EASEQM-4;	3=very easy 2=somewhat easy 1=not that easy 0=difficult
WKAINDEX	Walking ability index	Summary measure of self-reported walking ability.	WKAINDEX=EASEQM + EASE1M	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)
EASE1F	Ease climbing 1 flight	Recode of BLDW10EZ where 6 is easiest; imputations applied for missing BLDW10EZ	EASE1F=7-BLDW10EZ; If BLDW10YN=1 then EASE1F=4-BLDIF	If BLDW10YN=9 then EASE1F=.; If BLDW10EZ=8 then BLDW10EZ=2; If BLDIF=8 then BLDIF=2; If BLDW10EZ<0 then EASE1F=7-BLDW20EZ; If BLDW10YN=1 and (BLDIF=8 or BLDIF<0) then EASE1F=2; If BLDW10YN=0 and BLDW10EZ<0 and BLDW10EZ<0 then EASE1F=4; if EASE1F=. and BLDW10YN>9 and BLDW20YN>0 and (BLDW20YN=1 or BLDW20YN=8 or BLDW20YN†.Z) THEN EASE1F=4; if EASE1F=. and BLDW20YN=0 and BLDW20EZ†.Z then EASE1F=5	6=very easy 5=somewhat easy 4=not that easy 3=a little difficult 2=somewhat difficult 1=very difficult 0=unable to do
EASE2F	Ease climbing 2 flights	Recode of BLDW20EZ where 3 is easiest; imputations applied for missing BLDW20EZ	If BLDW20YN=1 then EASE2F=0; if BLDW20EZ=3 then EASE2F=1; if BLDW20EZ=2 then EASE2F=2; if BLDW20EZ=1 then EASE2F=3	If (BLDW20YN=8 or BLDW20YN<0) and (BLDW20EZ<0 or BLDW20EZ=8) and EASE1F>0 then EASE2F=1; if BLDW20YN†.Z and BLDW20EZ†.Z then EASE2F=EASE1F-4; if BLDW20YN=0 and (BLDW20EZ†.Z or BLDW20EZ=8) and EASE1F>0 then EASE2F=EASE1F-4; if BLDW10YN=1 then EASE2F=0; if EASE1F-3 < EASE2F and EASE1F†3 then EASE2F=0, else If EASE1F-3 < EASE2F then EASE2F=EASE1F-3	3=very easy 2=somewhat easy 1=not that easy 0=difficult
CSAINDEX	Climbing stairs ability index	Summary measure of self-reported walking ability.	CSAINDEX=EASE1F + EASE2F	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)

Documentation for Year 4 Calculated Variable Dataset
(Y4Calc)

YEAR 4 CALCULATED VARIABLE DATASET (Y4Calc)..... 33

Appendix I Calculated Variable List

Appendix II 20-Meter Walk Calculated Variables

Appendix III Ankle-Arm and Sitting Blood Pressure Calculated Variables

Appendix IV Anthropometric Measures -- Derived Variables

Appendix V Depression Scales

Appendix VI ECG Calculated Variables

Appendix VII Knee Pain Indicator Variables

Appendix VIII Long Distance Corridor Walk Calculated Variables

Appendix IX Performance Measure Calculated Variables

Appendix X Physical Activity Calculated Variables

Appendix XI Rose Scales

Appendix XII Self-Reported Function Calculated Variables

YEAR 4 CALCULATED VARIABLE DATASET (Y4Calc)

1. General description

The Y4calc contains 100 variables derived (calculated) as described below. The documentation is grouped by the type of data from which the variables are calculated (anthropometry, physical activity, etc.), and the variables are positioned in the dataset in the same groupings.

There are 3075 observations in the Y4calc file.

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under Y4calc) and in Appendix I. NOTE: SOME CALCULATED VARIABLES USE THE SAME VARIABLE NAMES EVERY YEAR. THE ANALYST IS REMINDED TO RENAME VARIABLES BEFORE MERGING WITH OTHER YEARS AS NECESSARY.

3. Dataset structure and contents

The Y4calc file contains a single observation per participant.

Key variables:

HABCID	Health ABC Enrollment ID# without the 2-letter prefix
GENDER	Gender (1=Male; 2=Female)*
RACE	Race (1=White; 2=Black)*
RACEGEN	Race-gender group (1=White male, 2=White female, 3=Black male, 4=Black female)*
SITE	Clinic Site (1=Memphis, 2=Pittsburgh)*

4. Condition of data

a. Strengths and weaknesses of dataset items: When a calculated variable includes imputations for missing variables, the method of imputation is **bolded** in the documentation.

Most importantly, the ECG raw data were found to have been previously imported with a column shift, causing extremely anomalous values for the PFT calculated variables. This import has been corrected (see Y4Read), and the ECG calculated variables (see Appendix VI) have been recomputed.

Two new variables have been added for this release: Y4WTK gives the year 4 weight in kg for all participants, regardless if they were measured in kg (in the clinic) or lb (in the home). This variable was added to allow easy comparison to other years. And the variable CES_D10, which is created in other years to allow longitudinal comparison of

* Must link to Health ABC participant history file (PH) to add this variable.

depression even when the short form of the Center for Epidemiologic Studies Depression scale is used, has been added.

As of this release, a thorough comparison across years of heights and weights has been carried out, which allowed us to identify a small number of impossible values. Where these could be traced to incorrectly scanned or edited data, the data were corrected. When it was impossible to know what the correct value was, but it was certain that the recorded value was incorrect (e.g., a participant's height having changed by 6 inches in one year and then changed back the next time it was measured), the erroneous data were simply blanked out. Thus six values of D2SH, 15 values of BMI and 5 of BMICAT have changed.

For the CES_D score (full 20-item scale), a maximum of 4 missing or refused items are imputed from the non-missing values. For CES_D10 (the short-form 10-item scale), a maximum of 1 missing value was previously imputed. After consultation with the prime mover, this maximum was changed to 2 for the CES_D10, causing two participants to have scores who previously did not.

A bug in the Year 4 self-reported function calculated variable (see Appendix XII) code for EASEUP has been corrected, resulting in the ceiling for this variable being raised from 5 to 6 to match other years. Consequently, 1775 values of this variable changed, some from missing to a non-missing value, and others from a value of 5 to a value of 6.

The Long Distance Corridor Walk (LDCW) calculated variables (see Appendix VIII) have been corrected to handle proxy clinic visits properly. Previously, EXCLUDE4 was erroneously set to 9 (No clinic visit); HOWFARY4, Y4DID2MW, Y4DID400 were set to 0; and Y4V2MINM, Y4V400MT, Y4400MSD, and TWOMINSD were set to missing (.). Now, if the LDCW was done for a proxy clinic visit, the data are included.

The performance calculated variables (see Appendix IX) have been thoroughly examined and corrected. The investigator who designed these variables created them to mimic the way things were done in the EPESE study, so certain recodes were made when calculating them. Most importantly, in EPESE, the full-tandem stand was not supposed to be administered unless the semi-tandem stand was held for 10 seconds. In Health ABC the full-tandem stand was supposed to be administered if the ST stand has held for least 1 second. Similarly, in EPESE, the one-leg stand was not administered if the full-tandem stand was not held for the full 30 seconds, while in Health ABC this stand was administered if the participant held the full-tandem for at least 1 second. In addition, balance walks variables originally set to 0 when refused or entirely missing are now set to missing. These corrections resulted in a number of changes to these variables, and analysts are advised to rerun any analyses pertaining to these variables using the latest dataset.

An additional three variables have been renamed. SIXMPACE has been renamed Y4UWPACE, SIXRATIO has been renamed Y4UWRATIO, and UWSCORE has been renamed Y4UWSCR, and these variables have been broadened to include performance on the 3- or 4-meter walk administered in the home in place of the 6-meter walk

administered in the clinic. Since each of these variables is standardized to the distance walked, their interpretation should be equivalent. The variables DID6MUW and SIXMWTM, which are not standardized to distance walked, have not been renamed or expanded.

Examiners were somewhat inconsistent in their recording of the ankle-arm blood pressure measurements when there was difficulty locating the tibial pulse. In most cases when “unable to locate tibial artery” was marked, the corresponding tibial systolic measurements were missing. However, in some cases, this value was completed as 0. Similarly, there were several cases where the tibial systolic measurements were marked 0, but only a few of them also had “unable to locate tibial artery” marked. The tibial systolic measurements should only be 0 if the tibial artery pulse was located before cuff inflation and then never heard again after complete deflation of the cuff.

Some of the physical activity questions were asked differently depending on whether the participant had a home or clinic visit, so the detailed documentation in Appendices VII, IX, X, and XII are separated by source of data (clinic vs. core home visit). Thus, a variable might be missing for one subset of participants (usually those who had a home visit, e.g. EASEUP and EASEHHW), but not for the other. In some cases, the summary variable calculated from individual calculated variables could be calculated when the individual calculated variable could not. This is particularly the case for WSKKWK, because the composite variable DATWKKWK could be substituted for the missing individual variables DAEWKKWK and DAOWKKWK, which could not be created for either dataset in year 4. The variables WALKTIME, WALKCAT, and WSKKWK, which can be calculated for home visits but weren’t previously, have been added as of release version 4.011.

In addition, Y4RKP30D and Y4LKP30D have been set to missing for participants asked the home visit version of the question because this version of the question cannot be used to reliably assign the knee pain to the left or right knee. Y4RWOMAC and Y4LWOMAC have also been set to missing for home visit participants (Y4LPLIK and Y4RPLIK are calculated for these participants, see documentation in Appendix VII for clarification of the difference between these two sets of variables).

5. Dataset index formulation and key variable mapping

The Y4calc file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

Appendix I
CALCULATED VARIABLE LIST
(LINKED)

Variable	Variable Description	Grouping
Y4R20MSD	fast walking speed (m/sec) over 20m	20-meter walk
Y4U20MSD	usual walking speed (m/sec) over 20m	20-meter walk
AAIL	Ankle-arm index, left leg	Ankle-arm and Sitting BP Calc Vars
AAIR	Ankle-arm index, right leg	Ankle-arm and Sitting BP Calc Vars
DIABP	Avg sitting diastolic BP (mm Hg)	Ankle-arm and Sitting BP Calc Vars
LOWAAI	Lower extremity arterial disease index	Ankle-arm and Sitting BP Calc Vars
MINAAI	Lowest ankle-arm index, right or left	Ankle-arm and Sitting BP Calc Vars
SYSBP	Avg sitting systolic BP (mm Hg)	Ankle-arm and Sitting BP Calc Vars
BMI	Body Mass Index (kg/m ²)	Anthropometric Measures
BMICAT	Body Mass Index Category	Anthropometric Measures
Y4WTK	Weight (kg)	Anthropometric Measures
D2SH	Average of the Y4 standing height measurements (mm)	Anthropometric Measures
CES_D	CES-D Score	CES-D Score
CES_D10	CES_D short form score (clinic visit)	CES-D Score
ABHR	Abnormal heartrate	ECG Derived Variables
ARYHTTYP	Type of arrhythmia code	ECG Derived Variables
AVDEFTYP	Type of atrioventricular conduction defect	ECG Derived Variables
AXISAB	Axis abnormality	ECG Derived Variables
LONGPR	Long P-R interval	ECG Derived Variables
LONGQRS	Long QRS duration	ECG Derived Variables
LONGQT	Long QT duration	ECG Derived Variables
QMI	Q-wave myocardial infarction	ECG Derived Variables
QWVAB	Q-wave myocardial infarction	ECG Derived Variables
STSEGDEP	ST segment depression	ECG Derived Variables

Variable	Variable Description	Grouping
STSEGELV	ST segment elevation	ECG Derived Variables
STWVAB	ST or T wave abnormality	ECG Derived Variables
TECHFLAG	Flag for existence of technical problem interfering with ECG coding	ECG Derived Variables
TWVITMS	T-wave items	ECG Derived Variables
VDEF COD	Ventricular conduction defect code	ECG Derived Variables
Y4KP12MO	Knee pain (either knee) most days past 12 months	Knee pain
Y4KP30DA	Knee pain (either) most days in past 30	Knee pain
Y4KPACT	Knee pain with activity (either knee)	Knee pain
Y4KPSIG	Knee pain case - either knee	Knee pain
Y4LKP12M	Left knee pain most days past 12 months	Knee pain
Y4LKP30D	Left knee pain most days in past 30	Knee pain
Y4LKPACT	Left knee pain with activity	Knee pain
Y4LKPSIG	Left knee pain case	Knee pain
Y4LPLIK	Left knee activity pain - Likert scale (home visit only)	Knee pain
Y4LWOMAC	Left knee activity pain - Likert scale (clinic visit only)	Knee pain
Y4RKP12M	Right knee pain most days past 12 months	Knee pain
Y4RKP30D	Right knee pain most days in past 30	Knee pain
Y4RKPACT	Right knee pain with activity	Knee pain
Y4RKPSIG	Right knee pain case	Knee pain
Y4RPLIK	Right knee activity pain - Likert scale (home visit only)	Knee pain
Y4RWOMAC	Right knee activity pain - Likert scale (clinic visit only)	Knee pain

Variable	Variable Description	Grouping
D32SUM	2 minute distance	Long Distance Corridor Walk Calculated Vars.
D34SUM	400M distance	Long Distance Corridor Walk Calculated Vars.
EXCLUDE4	LDCW exclusion code	Long Distance Corridor Walk Calculated Vars.
HOWFARY4	LDCW completion status	Long Distance Corridor Walk Calculated Vars.
Y4400MSD	walking speed (m/sec) over 400 meters	Long Distance Corridor Walk Calculated Vars.
Y4DID2MW	completed 2 minute walk yes/no	Long Distance Corridor Walk Calculated Vars.
Y4DID400	completed 400m walk yes/no	Long Distance Corridor Walk Calculated Vars.
Y4V2MINM	meters walked in 2 min - complete only	Long Distance Corridor Walk Calculated Vars.
Y4V400MT	time to walk 400m - complete only	Long Distance Corridor Walk Calculated Vars.
ABLE5CS	Did 5 chair stands	Performance Measure Calculated Vars.
CAT5CS	EPESSE score for chair stands	Performance Measure Calculated Vars.
CHR5PACE	Chair stands per second	Performance Measure Calculated Vars.
CSRATIO	Chair stands performance ratio	Performance Measure Calculated Vars.
DID6MNW	Did 6m narrow walk	Performance Measure Calculated Vars.
DID6MUW	Did 6m usual walk	Performance Measure Calculated Vars.
EPESPPB	EPESSE performance battery score 0-12	Performance Measure Calculated Vars.
FSBRATIO	Standing balance performance ratio	Performance Measure Calculated Vars.
FSBTIME	Full standing balance test time 0-90	Performance Measure Calculated Vars.
HABCPPB	Health ABC performance score 0-4	Performance Measure Calculated Vars.
NWPACE	Walking speed for narrow walk 6m	Performance Measure Calculated Vars.

Variable	Variable Description	Grouping
NWRATIO	Narrow walk performance ratio	Performance Measure Calculated Vars.
NWSCORE	EPESE category for narrow walk	Performance Measure Calculated Vars.
NWSCOREQ	% diff btw narrow & usual walks	Performance Measure Calculated Vars.
NWTIME	Time to walk a 20cm wide 6m course	Performance Measure Calculated Vars.
PPB12CAT	Categorical scoring of HABCPPB 0-12	Performance Measure Calculated Vars.
SBSCORE	EPESE score for standing balance	Performance Measure Calculated Vars.
SIXMWTM	Time to walk 6m	Performance Measure Calculated Vars.
Y4UWPACE	Walking speed (m/sec) over 3,4, or 6m	Performance Measure Calculated Vars.
Y4UWRATIO	Usual walk performance ratio	Performance Measure Calculated Vars.
Y4UWSCR	EPESE categories for walking speed	Performance Measure Calculated Vars.
BKTWTIME	minutes walking briskly/week	Physical Activity Calculated Vars
BRISK180	walks briskly \geq 180 min/week	Physical Activity Calculated Vars
BRISK90	walks briskly \geq 90 min/week	Physical Activity Calculated Vars
DACWKKWK	kcal/kg/week doing child/adult care	Physical Activity Calculated Vars
DAFSKKWK	kcal/kg/week climbing stairs	Physical Activity Calculated Vars
DAMCKKWK	Kcal/ kg/ week doing major chores	Physical Activity Calculated Vars
DAPWKKWK	kcal/kg/week doing paid work	Physical Activity Calculated Vars
DATWKKWK	Kcal/ kg/ week total walking	Physical Activity Calculated Vars
DAVWKKWK	kcal/kg/week doing volunteer work	Physical Activity Calculated Vars
WALKCAT	minutes walking/week category	Physical Activity Calculated Vars

Variable	Variable Description	Grouping
WALKTIME	minutes walking/week	Physical Activity Calculated Vars
WSKKWK	kcal/kg/week - walking + stairs	Physical Activity Calculated Vars
WVCKKWK	kcal/kg/week - work, vol, caregiving	Physical Activity Calculated Vars
ROSEANG	Rose angina score	Rose scales
ROSEIC	Rose intermittent claudication score	Rose scales
CSAINDEX	Climbing stairs ability index	Self-Reported Function Calculated Vars.
EASE10P	Ease lift/carry 10 lbs	Self-Reported Function Calculated Vars.
EASE1F	Ease climbing 1 flight	Self-Reported Function Calculated Vars.
EASE1M	Ease walking 1 mile	Self-Reported Function Calculated Vars.
EASE20P	Ease lift/carry 20 lbs	Self-Reported Function Calculated Vars.
EASE2F	Ease climbing 2 flights	Self-Reported Function Calculated Vars.
EASEQM	Ease walking 1/4 mile	Self-Reported Function Calculated Vars.
EASEUP	Ease rising from chair	Self-Reported Function Calculated Vars.
LCAINDEX	Lift/carry ability index	Self-Reported Function Calculated Vars.
WKAINDEX	Walking ability index	Self-Reported Function Calculated Vars.

Appendix II
20-Meter Walk Calculated Variables

Investigator Name: Eleanor Simonsick

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4R20MSD	Fast walking speed (m/sec) over 20m	Walking speed (m/sec) during the fast-paced 20 meter walk	$Y4R20MSD = 20 / D320TM2S$	If $D320TM2S = 0^*$, then $Y4R20MSD = .M$; if $D320TM2S \leq .z$, then $Y4R20MSD = .A$	m/sec
Y4U20MSD	Usual walking speed (m/sec) over 20m	Walking speed (m/sec) during the usual pace 20 meter walk	$Y4U20MSD = 20 / D320TM1S$	If $D320TM1S = 0^*$, then $Y4U20MSD = .M$; if $D320TM1S \leq .z$, then $Y4U20MSD = .A$	m/sec

*When a participant tried but was unable to complete either the usual pace or fast pace walk, both the time variable ($D320TM1s$ or $D320TM1S$, respectively) and the number of steps variable ($D320STP1$ or $D320STP2$, respectively) were coded as 0.

```

/*****
*****
* Program Name: 20mw_y4.sas (SAS 9.1)
* Saved as: \\Fu-hsing-c\habc\habc_sas\Calculated Variables\Programs\Year
4\20mw_y4.sas
* Study: HABC
* Purpose: Creates calculated variables for Year 4 for the 20 meter walk
* Input: habc4.d3
* Output: calc.ldcw
* Programmer: Pat Spurr
* Date: Existing program modified 8/14/06 by Pat Spurr
*****
*****/
/*
proc datasets library=work memtype=data kill; quit;
*/

%include 'i:\habc_sas\macros\remotesub.sas';
%libname calc 'e:\lakin\habc_sas\calculated variables\datasets\year 6';

%include '\\Fu-hsing-c\habc\habc_sas\programs\initv8.sas';

options ls=90 ps=56 formchar='|----+|----+=|-\<>*' nodate nofmterr;
title;

*****;
** 20 meter walk **;
*****;

* Added alternative code for source data (from "current" lib 9/24/02 EK);

data calc.ldcw(keep=habcid y4u20msd y4r20msd);
  *set current.y4clnvis(keep=habcid d320tmls d320tm2s);
  set habc4.d3(keep=habcid d320tmls d320tm2s);

  if d320tmls>0 then Y4U20MSD=20/d320tmls;
  else if d320tmls=0 then y4u20msd=.M;
  else y4u20msd=.A;

  if d320tm2s>0 then Y4R20MSD=20/d320tm2s;
  else if d320tm2s=0 then y4r20msd=.M;
  else y4r20msd=.A;

label y4u20msd='Usual walking speed over 20 meters'
      y4r20msd='Rapid walking speed over 20 meters';
run;

/*
* Check for duplicates;
data chk;
  set calc.ldcw;
  by habcid;
  if not (first.habcid and last.habcid);
run;
proc print data=chk;
title4 'Duplicates in ldcw Y4';
run;

```

```
title4;

proc contents data=calc.ldcw;
run;
proc univariate data=calc.ldcw;
    var y4r20msd y4u20msd;
run;
*/
```

Appendix III
Ankle-Arm and Sitting Blood Pressure Calculated Variables

Investigator Name: Mikki Danielson

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
AAIR	Ankle-arm index, right leg	Average of two measurements of ratio between right tibial and right brachial systolic blood pressure	AAI1R=D2AARP1/D2AARB1 AAI2R=D2AARD2/D2AARB2 AAIR=(AAI1R+AAI2R)/2	If D2AAPR = 0 then missing If one trial missing, use single measure for calculation	unitless
AAIL	Ankle-arm index, left leg	Average of two measurements of ratio between left tibial and right brachial systolic blood pressure	AAI1L=D2AALP1/D2AARB1 AAI2L=D2AALD2/D2AARB2 AAIL=(AAI1L+AAI2L)/2	If D2AAPR = 0 then missing If one trial missing, use single measure for calculation	unitless
MINAAI	Lowest ankle-arm index, right or left	Lower of left and right AAI, each measured twice and averaged)	MINAAI=lowest of either AAIR or AAIL	If right or left leg missing, use whichever leg is available	unitless
LOWAAI	Lower extremity arterial disease index	Lower extremity arterial disease group. Participants are categorized according to lowest AAI of either right or left w/ cutpoint of <0.90	LowAAI = 1 if MinAAI <0.90, 0 if MinAAI >=0.90	If D2AAPR=0 then missing	unitless

Documentation date: 2010-10-01

Ankle-Arm and Sitting Blood Pressure Calculated Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
SYSBP	Avg sitting systolic BP	Avg of two measurements of sitting systolic BP (mmHg)	$SYSBP=D2SYS$ or $SYSBP=(Z4SYS+Z4SY2)/2$ or $SYSBP=(YCSYS+YCSY2)/2$	If only 1 trial, then use that trial as BP	mm Hg
DIABP	Avg sitting diastolic BP	Avg of two measurements of sitting diastolic BP (mmHg)	$DIABP=D2DIA$ or $SYSBP=(Z4DIA+Z4DIA2)/2$ or $SYSBP=(YCDIA+YCDIA2)/2$	If only 1 trial, then use that trial as BP; If D2DIA or D2DIA2<30, set them =30 before averaging	mm Hg

```

/*FILENAME EXEC 'C:\SAS\AUTOEXEC.SAS' ;
%INCLUDE EXEC ;
SIGNON MISFIT;
RSUBMIT ;*/

/*%include 'i:\habc_sas\programs\initv8.sas';*/

*****
Updated from year 1 code
** Calculated variables requested by Michael Nevett, in email rec'd **
** from Emily Kenyon, dated 4/5/00 **
** ** **
** Fran Harris 4/6/00 **
*****;
/*libname daf 'e:\macuam\habc_sas\data analysis file';
filename intr 'e:\macuam\habc_sas\programs\intr.sas';*/
/*libname calc 'e:\macuam\habc_sas\calculated variables\datasets\Year 4';*/
/*%include intr; *contains libname/filename;
options ls=132 ps=58 formchar='|----|+|----+=|-\<>*' nocenter pageno=1 nofmterr;
title 'HABC calculated variables';
title2 'Program: bp.calcvar.sas';*/
*added alternative code for source data (from "current" lib 9/24/02 EK);
/*libname calc 'i:\habc_sas\calculated variables\datasets\Year 4';*/
data bph(keep=habcid sysbp diabp);
  set habc3.z4(keep=habcid z4sys z4sy2 z4dia z4dia2 z4type);
  if z4type=4;
*set current.y4corehv(keep=habcid z4sys z4sy2 z4dia z4dia2);
  if z4dia >.z then z4dia =max(z4dia ,30);
  if z4dia2>.z then z4dia2=max(z4dia2,30);
  SYSBP=(z4SYS+z4SY2)/2;
  DIABP=(z4DIA+z4DIA2)/2;
  IF z4SYS<=.z THEN SYSBP=z4SY2; ELSE
  IF z4SY2<=.z THEN SYSBP=z4SYS;
  IF z4DIA<=.z THEN DIABP=z4DIA2; ELSE
  IF z4DIA2<=.z THEN DIABP=z4DIA;
run;

data bpp(keep=habcid sysbp diabp);
  set habcff.yc(keep=habcid ycsys ycsy2 ycdia ycdia2 ycvisit);
  if ycvisit=4;
*set current.y4proxy (keep=habcid ycsys ycsy2 ycdia ycdia2);
  if ycdia >.z then ycdia =max(ycdia ,30);
  if ycdia2>.z then ycdia2=max(ycdia2,30);
  SYSBP=(ycSYS+ycSY2)/2;
  DIABP=(ycDIA+ycDIA2)/2;
  IF ycSYS<=.z THEN SYSBP=ycSY2; ELSE
  IF ycSY2<=.z THEN SYSBP=ycSYS;
  IF ycDIA<=.z THEN DIABP=ycDIA2; ELSE
  IF ycDIA2<=.z THEN DIABP=ycDIA;
run;

data bpc;
  set habc4.d2(keep=habcid d2sys d2dia
                d2aarp1 d2aarp2
                d2aarb1 d2aarb2
                d2aapr
                d2aalp1 d2aalp2);
*set current.y4clnvis (keep=habcid d2sys d2dia
                      d2aarp1 d2aarp2

```

```

                                d2aarb1 d2aarb2
                                d2aapr
                                d2aalp1 d2aalp2);

by habcid;
sysbp=d2sys;
diabp=d2dia;
**corrected 11/19/02 to set all diabp<30 to 30;
if diabp>.z and diabp<30 then diabp=30;
AAI1R=d2AARP1/d2AARB1;
AAI2R=d2AARP2/d2AARB2;
if d2AAPR=0 then do;
    AAI1R=.;
    AAI2R=.;
END;
AAIR=(AAI1R+AAI2R)/2;
if AAI1R=. THEN AAIR=AAI2R; ELSE
IF AAI2R=. THEN AAIR=AAI1R;

AAI1L=d2AALP1/d2AARB1;
AAI2L=d2AALP2/d2AARB2;
if d2AAPR=0 then do;
    AAI1L=.;
    AAI2L=.;
END;
AAIL=(AAI1L+AAI2L)/2;
if AAI1L=. THEN AAIL=AAI2L; ELSE
IF AAI2L=. THEN AAIL=AAI1L;

MINAAI=MIN(AAIR, AAIL);

if .z<MINAAI<0.90 THEN LOWAAI=1; ELSE
if MINAAI>=0.90 THEN LOWAAI=0;
RUN;

proc format;
    value lowaaif 1='Lowest AAI<0.90' 0='Lowest AAI>=0.90';
run;

data bp;
    set bph bpc bpp;
    by habcid;
    if habcid=5097 and d2sys<0 then delete;
    *if not(first.habcid and last.habcid) then put 'DUPLICATE: ' _all_;
run;
/*
data junk;
    set bp;
    by habcid;
    if not(first.habcid and last.habcid);
run;

proc print data=junk;
    var habcid;
run;

proc print data=bp;
    where habcid=5097;
run;

```

```

proc print data=bp(obs=40);
  var AAI1R d2AARP1 d2AARB1 AAI2R d2AARP2 d2AARB2 d2AAPR aair;
run;
proc print data=bp(obs=40);
  var AAI1L d2AALP1 d2AARB1 AAI2L d2AALP2 d2AARB2 d2AAPR aail;
run;
proc print data=bp(obs=40);
  var AAIR AAIL MINAAI LOWAAI;
run;
proc print data=bp(obs=40);
  where d2aarb1<0 or d2aarb2<0 or d2aarp1<0 or d2aarp2<0;
  var aair d2aarp1 d2aarb1 aai2r d2aarp2 d2aarb2 aair;
run;
proc print data=bp(obs=40);
  where d2aarb1<0 or d2aalp1<0 or d2aalp2<0 or d2aarb2<0;
  var aail d2aalp1 d2aarb1 aai2l d2aalp2 d2aarb2 aail;
run;
proc print data=bp;
  where aair<0 or aail<0;
  var minaaai aair aail;
run;
proc print data=bp;
  where d2aapr=0;
  var lowaaai minaaai;
run;

proc print data=bp;
  where minaaai=0;
  var minaaai aail aai2l aai2r d2aalp1 d2aarb1 d2aalp2 d2aarb2;
run;*/

proc format;
  value lowaaif 1='Lowest AAI<0.90' 0='Lowest AAI>=0.90';
run;
data calc.y4bp;
  set bp(keep=habcid aair aail minaaai lowaaai sysbp diabp);
  label aair='Ankle-arm index, rt leg'
        aail='Ankle-arm index, lt leg'
        minaaai='Lowest ankle-arm index'
        lowaaai='Lower extr arterial dz index'
        sysbp='Avg sitting systolic BP, mm Hg'
        diabp='Avg sitting diastolic BP, mm Hg';
  format lowaaai lowaaif.;
run;
/*proc univariate data=calc.y4bp;
  var aair aail minaaai;
run;
proc sort data=calc.y4bp out=bp;
  by lowaaai;
run;
proc univariate data=bp;
  by lowaaai;
  var minaaai;
run;
proc contents data=calc.y4bp;
run; */

```

Appendix IV
Anthropometric Measures -- Derived Variables

Investigator Name: Jill A. Bennett
Analysis Plan Reference #: AP99-36

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
BMI	Body Mass Index	Weight in kg / height in meters squared	1. Weight in kg = D2WTK or YCWT (if YCLBSKG=2) or 0.45359*Z4WTLBS or 0.45359*YCWY (if YCLBSKG=1) 2. Height in meters squared = D2SH divided by 1000 (to change mm to meters) then that number is squared 3. Divide weight in kg by the number derived in step 2	If D2WTK, Z2WTLBS, and YCWT are all missing, BMI=.M If D2SH is missing, use P2SH (Year 1 height)	kg/m ²
Y4WTK	Weight	Weight in kg	Y4WTK = D2WTK or YCWT (if YCLBSKG=2) or 0.45359*Z4WTLBS or 0.45359*YCWY (if YCLBSKG=1)	n/a	kg

Investigator Name: Michael Nevitt

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
BMICAT	Body Mass Index Category	Body mass categories according to NIH guidelines	BMICAT=1 if BMI<25 BMICAT=2 if 25≤BMI<30 BMICAT=3 if BMI ≥ 30	If BMI is missing, BMICAT=.M	1=normal 2=overweight 3=obese

Anthropometric Measures -- Derived Variables

Investigator Name: Emily Kenyon and Susan Rubin

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
D2SH	Average of the baseline standing height measurements	Average of 1-4 baseline standing height measurements	If only 2 standing height measurements were obtained (D2SH1 and D2SH2), take an average of D2SH1 and D2SH2. If all 4 measurements were made, average D2SH3 and D2SH4 only. If for some reason 3 measurements were made -- which is not according to protocol, average all 3 of these measurements. If only one measurement was obtained, use this as the baseline standing height measurement.	If D2SH1, D2SH2, D2SH3, and D2SH4 are all missing, D2SH=.M	mm

```

* YEAR 3 BMI - calculated from y4 height and y4 weight unless y4 height is
missing*;
/*%include 'i:\habc_sas\programs\initv8.sas'; *contains libname/filename;
libname current 'i:habc_sas\data analysis file\current';

libname calc 'i:\habc_sas\calculated variables\datasets\year 4'; */

*CALC: Temporary place for test datasets*;

%macro avgmeas(var);
nmissstot=nmiss(of &var.1-&var.4);
if nmissstot=0 then do;
&var=mean(&var.3,&var.4);
if max(&var.3,&var.4)>min(&var.3,&var.4)*1.2 then
put habcid '09'x "&var.1" '09'x &var.1 '09'x
&var.2" '09'x &var.2 '09'x
&var.3" '09'x &var.3 '09'x
&var.4" '09'x &var.4 '09'x ;
end; else
if nmissstot=1 then do;
&var=mean(&var.1,&var.2,&var.3,&var.4);
if max(&var.1,&var.2,&var.3,&var.4)>min(&var.1,&var.2,&var.3,&var.4)*1.2
then
put habcid '09'x "&var.1" '09'x &var.1 '09'x
&var.2" '09'x &var.2 '09'x
&var.3" '09'x &var.3 '09'x
&var.4" '09'x &var.4 '09'x ;
end; else
if nmissstot=2 then do;
&var=mean(&var.1,&var.2,&var.3,&var.4);
if max(&var.1,&var.2,&var.3,&var.4)>min(&var.1,&var.2,&var.3,&var.4)*1.2
then
put habcid '09'x "&var.1" '09'x &var.1 '09'x
&var.2" '09'x &var.2 '09'x
&var.3" '09'x &var.3 '09'x
&var.4" '09'x &var.4 '09'x ;
end; else
if nmissstot=3 then &var=max(&var.1,&var.2,&var.3,&var.4);
%mend avgmeas;

data ylhgt;
set current.y1calc(keep=habcid p2sh);
run;
*added alternative code for source data (from "current" lib 9/24/02 EK);
data d2;
set habc4.d2(keep=habcid d2wtk d2sh1 d2sh2 d2sh3 d2sh4);
*set current.y4clnvis (keep=habcid d2wtk d2sh1 d2sh2 d2sh3 d2sh4);
%avgmeas(D2SH);
run;
data z4;
set habc3.z4(keep=habcid z4wtlbs z4type);
if z4type=4;
z4wtk=0.45359*z4wtlbs;
* set current.y4corehv (keep=habcid z4wtk);
run;
data yc;

```

```

set habcff.yc(keep=habcid ycvisit ycwtk yclbskg);
if ycvisit=4;
if yclbskg=1 then do;
    ycwtk=ycwt*0.45359;end;
else do;
    ycwtk=ycwt; end;
*set current.y4proxy(keep=habcid ycwtk)
run;

data calc;
merge ylhgt(in=iny1) d2(in=iny4);
by habcid;
if iny4 and iny1;
bmi1=d2wtk/((p2sh/1000)**2);
bmi4=d2wtk/((d2sh/1000)**2);
BMI=bmi4;
if bmi4<0 then do;
bmi=bmi1; end;
run;

data calch;
merge ylhgt(in=iny1) z4(in=iny4);
by habcid;
if iny4 and iny1;
BMI=z4wtk/((p2sh/1000)**2);
*delete the guy who had both a home and clinic visit;
if habcid=5097 then delete;
run;

data calcp;
merge ylhgt(in=iny1) yc(in=iny4);
by habcid;
if iny4 and iny1;
BMI=ycwtk/((p2sh/1000)**2);
run;
*Update to keep Y4WTK as a calculated variable EK 6/1/09;
data calc.y4bmi(keep=HABCID BMI BMICAT D2SH Y4WTK);
*data y4bmi(keep=habcid bmi bmicat d2sh);
set calc (rename=(d2wtk=Y4WTK)) calch (rename=(z4wtk=Y4WTK)) calcp
(rename=(ycwtk=Y4WTK));
by habcid;
if not (first.habcid and last.habcid) then put 'DUPLICATE: ' _all_;
if .z< bmi< 25 then BMICAT=1; else
if 25<=bmi< 30 then bmicat=2; else
if bmi>=30 then bmicat=3;
label bmicat='BMI category'
bmi='BMI, kg/m2'
d2sh='average Y4 standing height meas (mm)'
Y4WTK='Year 4 weight (kg)';
format bmicat bmicf.;
run;

proc sort data=calc.y4bmi;
by habcid;
run;

proc sort data=calc.y4bmi out=temp;

```

```
*proc sort data=y4bmi out=temp;  
  by bmicat;  
run;  
proc univariate data=temp;  
  by bmicat;  
  var bmi;  
  title4 'Distribution of BMI within each BMICAT';  
run;
```

Appendix V
Depression Scales

Investigator Name: Ronald Shorr, MD, MS
Analysis Plan Reference Number: AP98-08

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
1. CES_D	CES-D score (Core Home Visit only)	CES-D score calculated according to Radloff, L.S. (1977). The CES-D scale: a self report Major Depressive Disorder scale for research in the general population. Applied Psychological Measurement, 1, 385-401.	<ol style="list-style-type: none"> 1. For the following variables, use the following to convert the score: 1(rarely)=3, 2(some of time)=2, 3(much of time)=1, 4(most of time)=0: DBFGOOD, DBFHOPE, DBFHAPPY, DBFENJOY (or ZCFGOOD, ZCFHOPE, ZCFHAPPY, ZCFENJOY) 2. For the remainder of the variables, use the following conversion 1(rarely)=0, 2(some of time)=1, 3(much of time)=2, 4(most of time)=3 3. Sum the score of the 20 items. (max=60, min=0) 	<p>If any item is answered 8(Don't know) or 7 (Refused), set that item to missing.</p> <p>For 1-4* missing items, assign the average score of the answered items to the missing items. If >4 items are missing, CES_D is missing.</p>	unitless

* In Year 1, we used the erroneous cutoff of 5. Thus two participants (HA1457 and HA1624) had baseline CES_D scores that are invalid and should be set to missing in any change analyses.

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
CES_D10	CES_D short form score	CES-D short form score calculated according to Andresen, E.M. and Malmgren, J.A. (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D	<ol style="list-style-type: none"> 1. For DBFHAPPY and DBFHOPE (or ZCFHAPPY and ZCFHOPE) use the following to convert the score: 1(rarely)=3, 2(some of time)=2, 3(much of time)=1, 4(most of time)=0: 2. For the remainder of the variables, use the following conversion 1(rarely)=0, 2(some of time)=1, 3(much of time)=2, 4(most of time)=3 3. Sum the score of the 10 items. (max=30, min=0) 	<p>If any item is answered 8 (Don't know) or 7 (Refused), set that item to missing.</p> <p>For 1 missing item, assign the average score of the answered items to the missing items. If >2 items are missing, CES_D10 is missing.</p>	unitless

Analyst note: Although the range for the CES_D10 (short form) score is exactly half that of the CES_D (long form) score, there is no general agreement about how the scores compare. That is, there is not general agreement that the former can simply be doubled to make it analogous to the latter.

```

*****;
** Create CES-D (cesd) Scores for Core Home Visit **;
*****;
*****;
** Checks, corrections made to calculation of CES-D: **
** - Check that values >4 are set to missing (already done) **
** - Check that 4 vars reversed (ZCFGGOOD, ZCFHOPE, ZCFHAPPY, and **
** ZCFENJOY) (already done) **
** - Use scale of 0-3 instead of 1-4 (fixed here) **
** - Set CES-D to missing if >=5 scores missing (fixed here) **
** - Fill in avg score for missing scores if <= 5 missing (fixed here)**
**
** Fran Harris 4/4/00 **
** Updated for Year 4 by Emily Kenyon 9/2/02 **
**
** - Changed allowable number of missings from 1 to 2 for **
** CES_D10 score **
**
** Emily Scott 06/26/09 **
*****;
*added alternative code for source data (from "current" lib 9/24/02 EK);
data cesdh;
    set habc3.zc(keep=habcid zctype zccontac ZCfbothr ZCfeat ZCfblues ZCfgood
                ZCfmind ZCfdown ZCfeffrt Zcfhope
                ZCffail ZCffear ZCfsleep Zcfhappy
                ZCftalk ZCflone ZCfunfr ZCfenjoy
                ZCfcry ZCfsad ZCfdisme ZCfnogo);

    if zctype=4;
*set current.y4corehv (keep=habcid ZCfbothr ZCfeat ZCfblues ZCfgood
                    ZCfmind ZCfdown ZCfeffrt Zcfhope
                    ZCffail ZCffear ZCfsleep Zcfhappy
                    ZCftalk ZCflone ZCfunfr ZCfenjoy
                    ZCfcry ZCfsad ZCfdisme ZCfnogo);

    array c [20] c1-c20;
    array orig [20] ZCfbothr ZCfeat ZCfblues ZCfgood
                ZCfmind ZCfdown ZCfeffrt Zcfhope
                ZCffail ZCffear ZCfsleep Zcfhappy
                ZCftalk ZCflone ZCfunfr ZCfenjoy
                ZCfcry ZCfsad ZCfdisme ZCfnogo;

    do i=1 to 20;
        c[i]=orig[i];
        if c[i]>4 then c[i]=.M;
        if i in (4,8,12,16) then c[i]=5-c[i];
        c[i]=c[i]-1;
    end;
    avg_cesd=mean(of c1--c20);
    nmiss=nmiss(of c1--c20);
    if nmiss<=4 then do;
        if nmiss>0 then do i=1 to 20;
            if c[i]=. then c[i]=avg_cesd;
        end;
        CES_D=sum(of c1--c20);
    end;
    else if nmiss=20 and zccontac=2 then ces_d=.A;*Non-starred item on telephone
interview;
    else ces_d=.M;
    label ces_d='CES-D';

```

```

format ces_d 5.2;
run;

data cesdc;
  set habc4.db(keep=habcid dbfbothr dbfeat dbfblues dbfgood
              dbfmind dbfdown dbfeffrt dbfhope
              dbffail dbffear dbfsleep dbfhappy
              dbftalk dbflone dbfunfr dbfenjoy
              dbfcry dbfsad dbfdisme dbfnogo);
*set current.y4clnvis(keep=habcid dbfbothr dbfeat dbfblues dbfgood
                    dbfmind dbfdown dbfeffrt dbfhope
                    dbffail dbffear dbfsleep dbfhappy
                    dbftalk dbflone dbfunfr dbfenjoy
                    dbfcry dbfsad dbfdisme dbfnogo);

array c [20] c1-c20;
array orig [20] dbfbothr dbfeat dbfblues dbfgood
               dbfmind dbfdown dbfeffrt dbfhope
               dbffail dbffear dbfsleep dbfhappy
               dbftalk dbflone dbfunfr dbfenjoy
               dbfcry dbfsad dbfdisme dbfnogo;

do i=1 to 20;
  c[i]=orig[i];
  if c[i]>4 then c[i]=.M;
  if i in (4,8,12,16) then c[i]=5-c[i];
  c[i]=c[i]-1;
end;
avg_cesd=mean(of c1--c20);
nmiss=nmiss(of c1--c20);
if nmiss<=4 then do;
  if nmiss>0 then do i=1 to 20;
    if c[i]=. then c[i]=avg_cesd;
  end;
  CES_D=sum(of c1--c20);
end;
else ces_d=.M;
label ces_d='CES-D';
format ces_d 5.2;
run;

data cesd20 (keep=habcid ces_d);
  set cesdc cesdh;
run;
proc sort data=cesd20;
  by habcid;
run;

proc univariate data=cesd20;
  var ces_d;
run;
proc freq data=cesdh;
  tables nmiss
         ZCfbothr ZCfeat ZCfblues ZCfgood
         ZCfmind ZCfdown ZCfeffrt ZCfhope
         ZCffail ZCffear ZCfsleep ZCfhappy
         ZCftalk ZCflone ZCfunfr ZCfenjoy
         ZCfcry ZCfsad ZCfdisme ZCfnogo / list missing;
run;

```

```

proc print data=cesdh;
  where nmiss>4 and ces_d>0;
run;

proc freq data=cesdc;
  tables nmiss
    dbfbothr dbfeat dbfblues dbfgood
    dbfmind dbfdown dbfeffrt dbfhope
    dbffail dbffear dbfsleep dbfhappy
    dbftalk dbflone dbfunfr dbfenjoy
    dbfcry dbfsad dbfdisme dbfnogo / list missing;
run;

proc print data=cesdc;
  where nmiss>4 and ces_d>0;
run;

****Added 1/6/04: Create short version score for comparison to years with CES-
D-short;
data cesd10c;
  set habc4.db(keep=habcid dbfbothr dbfmind dbfdown dbfeffrt dbfhope
              dbffear dbfsleep dbfhappy dbflone dbfnogo);

  array c [10] c1-c10;
  array orig [10] dbfbothr dbfmind dbfdown dbfeffrt dbfhope
                dbffear dbfsleep dbfhappy dbflone dbfnogo;

  do i=1 to 10;
    c[i]=orig[i];
    if c[i]>4 then c[i]=.M;
    if i in (5,8) then c[i]=5-c[i];
    c[i]=c[i]-1;
  end;
  avg_cesd=mean(of c1-c10);
  nmiss=nmiss(of c1--c10);
  if nmiss<=2 then do;
    if nmiss>0 then do i=1 to 10;
      if c[i]=. then c[i]=avg_cesd;
    end;
    CES_D10=sum(of c1--c10);
  end;
  else ces_d10=.M;
  label ces_d10='CES-D 10';
  format ces_d10 5.2;
run;
proc univariate data=cesd10c;
  var ces_d10 avg_cesd;
run;
proc freq data=cesd10c;
  tables nmiss
    dbfbothr dbfmind dbfdown dbfeffrt dbfhope
    dbffear dbfsleep dbfhappy dbflone dbfnogo/ list missing;
run;

* and for home visits;
data cesd10h;
  set habc3.zc(keep=habcid zctype zccontac ZCfbothr ZCfmind ZCfdown ZCfeffrt
              ZCfhope);

```

```

                                ZCffear ZCfsleep ZCfhappy
                                ZCflone ZCfnogo);

    if zctype=4;
array c [10] c1-c10;
array orig [10] ZCfbothr
                ZCfmind ZCfdown ZCfeffrt ZCfhope
                ZCffear ZCfsleep ZCfhappy
                ZCflone ZCfnogo;

    do i=1 to 10;
        c[i]=orig[i];
        if c[i]>4 then c[i]=.M;
        if i in (5,8) then c[i]=5-c[i];
        c[i]=c[i]-1;
    end;
    avg_cesd=mean(of c1-c10);
    nmiss=nmiss(of c1--c10);
    if nmiss<=2 then do;
        if nmiss>0 then do i=1 to 10;
            if c[i]=. then c[i]=avg_cesd;
        end;
        CES_D10=sum(of c1--c10);
    end;
    else if nmiss=10 and zccontac=2 then ces_d10=.A;*Non-starred item on
telephone interview;
    else ces_d10=.M;
    label ces_d10='CES-D 10';
    format ces_d10 5.2;
run;

proc univariate data=cesd10h;
    var ces_d10 avg_cesd;
run;
proc freq data=cesd10h;
    tables nmiss
            ZCfbothr ZCfmind ZCfdown ZCfeffrt ZCfhope
            ZCffear ZCfsleep ZCfhappy ZCflone ZCfnogo/ list missing;
run;

data cesd10;
    set cesd10c(keep=habcid ces_d10) cesd10h(keep=habcid ces_d10);
run;
proc sort; by habcid;run;

data calc.cesd;
    merge cesd20 cesd10;
    by habcid;
run;

proc contents data=calc.cesd;
run;

```

Appendix VI
ECG Derived Variables

Investigator Name: Bernard R. Chaitman

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
TECHFLAG	Flag for existence of technical problem interfering with coding	If TECH_PRB=981 or 984, then flag=1. All others, flag=0	TECHFLAG=1 if TECH=981 or 984 Else, TECHFLAG=0	Missing=0	0=no serious problems 1=technical probs interfere coding
ARYTHTYP	Type of arrhythmia code	Combines ARRHYTH=31 and 32; 33 and 34; keeps 8 and 12 separate, and combines all other ARRHYTH codes	ARYTHTYP =0 if ARRHYTH is missing or 0 =1 if ARRHYTH =31 or 32 =2 if ARRHYTH =33 or 34 =3 if ARRHYTH =8 =4 if ARRHYTH =12 or 23 =5 if ARRHYTH =11, 13, 14, 15, 24, 41, 42, 51, 52, 61, 53, 64, 7, or 9 =6 if ARRHYTH =21 or 22	Missing=0	0=None 1=Persistent atrial fib. or flutter 2=Intermittent atrial fib. or flutter 3=Sinus bradycardia 4=Freq. premature vent. beats 5=Other 6=Persistent ventricular rhythm

ECG Derived Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
AVDEFTYP	Type of atrioventricular conduction defect	Combines AV_DEF=41 and 42; 22, 23, and 6. Keeps 1, 21, 3, and 5 separate.	AVDEFTYP =0 if AV_DEF is missing or 0 =1 if AV_DEF=1 =2 if AV_DEF=21 =3 if AV_DEF=3 =4 if AV_DEF=41 or 42 =5 if AV_DEF=5 =6 if AV_DEF=22, 23, or 6 =8 if AV_DEF=8	Missing=0	0=None 1=Complete AV block 2=Mobitz type II 3=Long P-R interval 4=persistent or intermittent WPW 5=short P-R interval 6=other 8=pacemaker
AXISAB	Axis abnormality	Categories for axis < -45 deg, -45≤axis≤120, axis>120 degrees	AXISAB =0 if -45≤AXIS≤120 =1 if AXIS<-45 =2 if AXIS>120	Missing if AXIS is missing	0=None 1=Left axis 2=Right axis
ABHR	Abnormal heartrate	Categories for normal HR, bradycardia (<50 bpm), tachycardia (>100 bpm)	ABHR =0 if 50≤HR≤100 =1 if HR<50 =2 if HR>100	Missing if HR is missing	0=No 1=Bradycardia 2=Tachcardia
LONGPR	Long P-R interval	Indicator variable for P-R interval > 220 msec	LONGPR =0 if PR ≤220 =1 if PR>220	Assign .T (missing due to technical problem) if blank	0=No 1=Yes

ECG Derived Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
QWVAB	Q-wave myocardial infarction	Major Q or QS abnormality for anterior, posterior, lateral or inferior	<p>MAJQWVAB =0 if Q_ANT, Q_POST, Q_LAT, and Q_INF are all missing or 0 or >39</p> <p>=1 if Q_ANT=11, 12, 16, 17, 21, 22, or 27; or if Q_INF=11, 12, 14, 15, 21, 22, 23, 24, or 25; or if Q_LAT=11, 12, 13, 21, 22, or 23; or if Q_POST=1</p> <p>=2 if Q_ANT=28, 31, or 32; or if Q_INF=26, 31, 34, 35, or 36; or if Q_LAT=28, 31, or 33; or if Q_POST=2</p>	All missing=0	0=None 1=Major 2=Minor
STWVAB	ST or T wave abnormality	Major ST or T wave abnormality for anterior, posterior, lateral or inferior	<p>STWVAB =0 if STD_LAT, STD_INF, STD_ANT, STL_LAT, STL_INF, STL_ANT, T_LAT, T_INF, and T_ANT are all missing or 0</p> <p>=1 if STD_LAT≥11; or if STD_INF≥11; or if STD_ANT≥11; or if STL_LAT≥2; or if STL_INF≥2; or if STL_ANT≥2; or if T_LAT=1 or 2; or if T_INF=1 or 2; or if T_ANT=1, 2, 5 or 6</p> <p>=2 if STD_LAT=2, 3 or 4; or if STD_INF=2,3, or 4; or if STD_ANT=2,3, or 4; or if L_LAT=3 or 4; or if T_INF=3 or 4; or if T_ANT=3 or 4</p>	All missing=0	0=None 1=Major 2=Minor

ECG Derived Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
QMI	Q-wave myocardial infarction	Presence of major Q-wave abnormality or minor Q-wave abnormality in the presence of a major ST or T-wave abnormality	QMI =0 if QWVAB=0 and STWVAB=0 =1 if QWVAB=1 or (QWVAB=2 AND STWVAB=1) all others =0	N/A	0=No 1=Yes
LONGQRS	Long QRS duration	QRS duration longer than 120 msec	LONGQRS =0 if QRS≤120 msec =1 if QRS>120 msec	Missing if QRS is missing	0=No 1=Yes
LONGQT	Long QT duration	QT duration longer than 460 msec	LONGQT =0 if QT≤460 =1 if QT>460	Missing if QT is missing	0=No 1=Yes
STSEGDEP	ST segment depression	ST segment depression excluding Q-wave myocardial infarction	STSEGDEP =0 if STD_LAT and STD_INF and STD_ANT are all missing or 0 or if QMI=1 =1 if STD_LAT≥11; or if STD_INF≥11; or if STD_ANT≥11 =2 if STD_LAT=2, 3 or 4; or if STD_INF=2, 3, or 4; or if STD_ANT=2, 3, or 4	All missing=0	0=None 1=Major 2=Minor
STSEGELV	ST segment elevation	ST segment elevation excluding Q-wave myocardial infarction	STSEGELV =0 if STL_LAT, STL_INF, and STL_ANT are all missing or 0 or if QMI=1 =1 if STL_LAT≥2; or if STL_INF≥2; or if STL_ANT≥2	All missing=0	0=None 1=Major

ECG Derived Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
TWVITMS	T-wave items	T-wave items excluding Q-wave myocardial infarction	<p>TWVITMS =0 if T_LAT, T_INF and T_ANT are all missing or 0 or if QMI=1</p> <p>=1 if T_LAT=1 or 2; or if T_INF=1 or 2; or if T_ANT=1, 2, 5 or 6</p> <p>=2 if L_LAT=3 or 4; or if T_INF=3 or 4; or if T_ANT=3 or 4</p>	All missing=0	0=None 1=Major 2=Minor
VDEFCOD	Ventricular conduction defect code	Combines VENT_DEF=12, 22, 3, 5, and 6; keeps 11, 21, 4, 7 and 8 separate`	<p>VDEFCOD =0 if VENT_DEF is missing or 0 =1 if VENT_DEF=11 =2 if VENT_DEF=21 =3 if VENT_DEF=4 =4 if VENT_DEF=7 =5 if VENT_DEF=8 =6 if VENT_DEF=12, 22, 3, 5 or 6</p>	Missing=0	0=None 1=Complete LBBB 2=Complete RBBB 3=Intravent. Block 4=left ant. hemiblock (LAH) 5=LAH with complete or intermittent RBBB 6=Incomplete or intermittent

```

title 'HABC calculated variables';
title2 'Program: ECG.calcvar.sas';

*****;
*%include '\\Fu-hsing-c\habc\habc_sas\programs\initV8.sas';
*libname calc 'M:\My Documents\2005 and beyond\SAS
HABC\HabcComparisonsJune2009\ECG\CopiedFromNetworkProgramFolder\Y4';

data ph;
  set daf.ph(keep=habcid);
run;
data ecg;
  set read4.ecglab4n(keep=habcid tech_prb arrhyth av_def axis hr pr
                    q_ant q_post q_lat q_inf
                    std_lat std_inf std_ant
                    stl_lat stl_inf stl_ant
                    t_lat t_inf t_ant
                    qrs qt vent_def);

run;

proc format;
  value techf 0='0:No serious problems'
             1='1:Technical problems';
  value arythtf 0='0:None'
               1='1:Persistent afib or flutter'
               2='2:Intermittent afib or flutter'
               3='3:Sinus bradycardia'
               4='4:Freq premature vent beats'
               5='5:Other'
               6='6:Persistent ventricular rhythm';
  value avdeftf 0='0:None'
               1='1:Complete AV block'
               2='2:Mobitz type II'
               3='3:Long P-R interval'
               4='4:Persistent or intermittent WPW'
               5='5:Short P-R interval'
               6='6:Other'
               8='8:Pacemaker';
  value axisabf 0='0:None'
               1='1:Left axis'
               2='2:Right axis';
  value abhrf 0='0:No' 1='1:Bradycardia' 2='2:Tachcardia';
  value waveabf 0='0:None' 1='1:Major' 2='2:Minor';
  value vdefcdf 0='0:None'
               1='1:Complete LBBB'
               2='2:Complete RBBB'
               3='3:Intravent block'
               4='4:Left ant hemiblock'
               5='5:LAH, complete or intermittent RBBB'
               6='6:Incomplete or intermittent';
  value ynfmt 0='0:No' 1='1:Yes' .t='Missing due to technical problem';
run;

data ecg;
  merge ph ecg;
  by habcid;

```

```

if tech_prb in ('981','984') then TECHFLAG=1; else
                    techflag=0;

if arrhyth in (' ','0') then ARYHTYP=0; else
if arrhyth in ('31','32') then arythtyp=1; else
if arrhyth in ('33','34') then arythtyp=2; else
if arrhyth='8' then arythtyp=3; else
if arrhyth in ('12','23') then arythtyp=4; else
if arrhyth in ('11','13','14','15','24','41','42','51','52',
                '61','53','64','7','9') then arythtyp=5; else
if arrhyth in ('21','22') then arythtyp=6;

if av_def in (' ','0') then AVDEFTYP=0; else
if av_def='1' then avdeftyp=1; else
if av_def='21' then avdeftyp=2; else
if av_def='3' then avdeftyp=3; else
if av_def in ('41','42') then avdeftyp=4; else
if av_def='5' then avdeftyp=5; else
if av_def in ('22','23','6') then avdeftyp=6; else
if av_def='8' then avdeftyp=8;

if -45<=axis<=120 then AXISAB=0; else
if .z< axis< -45 then axisab=1; else
if axis> 120 then axisab=2;

if 50<=hr<=100 then ABHR=0; else
if .z< hr< 50 then abhr=1; else
if hr> 100 then abhr=2;

if .z< pr<=220 then LONGPR=0; else
if pr> 220 then longpr=1; else
                    longpr=.t;

if (q_ant in (' ','0') or q_ant >'39') and
(q_post in (' ','0') or q_post>'39') and
(q_lat in (' ','0') or q_lat >'39') and
(q_inf in (' ','0') or q_inf >'39') then QWVAB=0; else
if q_ant in ('11','12','16','17','21','22','27') or
q_inf in ('11','12','14','15','21','22','23','24','25') or
q_lat in ('11','12','13','21','22','23') or
q_post='1' then qwvab=1; else
if q_ant in ('28','31','32') or
q_inf in ('26','31','34','35','36') or
q_lat in ('28','31','33') or
q_post='2' then qwvab=2;

if std_lat in (' ','0') and std_inf in (' ','0') and std_ant in (' ','0') and
stl_lat in (' ','0') and stl_inf in (' ','0') and stl_ant in (' ','0') and
t_lat in (' ','0') and t_inf in (' ','0') and t_ant in (' ','0')
then STWVAB=0; else
if std_lat in ('11','12') or
std_inf in ('11','12') or
std_ant in ('11','12','13','14','15','16','17','18','19') or
stl_lat in ('2','21','22','23','24','25','26','27','28','29') or
stl_inf in ('2','21','22','23','24','25','26','27','28','29') or
stl_ant in ('2','21','22','23','24','25','26','27','28','29') or

```

```

t_lat in ('1','2') or
t_inf in ('1','2') or
t_ant in ('1','2','5','6') then stwvab=1; else
if std_lat in ('2','3','4') or
std_inf in ('2','3','4') or
std_ant in ('2','3','4') or
t_lat in ('3','4') or
t_inf in ('3','4') or
t_ant in ('3','4') then stwvab=2;

if qwvab=0 and stwvab=0 then QMI=0; else
if qwvab=1 or (qwvab=2 and stwvab=1) then qmi=1; else
qmi=0;

if .z<=qrs<=120 then LONGQRS=0; else
if qrs> 120 then longqrs=1;

if .z<=qt<=460 then LONGQT=0; else
if qt> 460 then longqt=1;

if (std_lat in (' ','0') and std_inf in (' ','0') and std_ant in (' ','0'))
or qmi=1 then STSEGDEP=0; else
if std_lat in ('11','12') or
std_inf in ('11','12') or
std_ant in ('11','12','13','14','15','16','17','18','19') then stsegdep=1;
else
if std_lat in ('2','3','4') or
std_inf in ('2','3','4') or
std_ant in ('2','3','4') then stsegdep=2;

if (stl_lat in (' ','0') and stl_inf in (' ','0') and stl_ant in (' ','0'))
or qmi=1 then STSEGELV=0; else
if stl_lat in ('2','21','22','23','24','25','26','27','28','29') or
stl_inf in ('2','21','22','23','24','25','26','27','28','29') or
stl_ant in ('2','21','22','23','24','25','26','27','28','29')
then stsegelv=1;

if (t_lat in (' ','0') and t_inf in (' ','0') and t_ant in (' ','0'))
or qmi=1 then TWVITMS=0; else
if t_lat in ('1','2') or
t_inf in ('1','2') or
t_ant in ('1','2','5','6') then twvitms=1; else
if t_lat in ('3','4') or
t_inf in ('3','4') or
t_ant in ('3','4') then twvitms=2;

if vent_def in (' ','0') then VDEF COD=0; else
if vent_def = '11' then vdefcod=1; else
if vent_def = '21' then vdefcod=2; else
if vent_def = '4' then vdefcod=3; else
if vent_def = '7' then vdefcod=4; else
if vent_def = '8' then vdefcod=5; else
if vent_def in ('12','22','3','5','6') then vdefcod=6;

label techflag='Technical problem interfering with coding'
arythtyp='Type of arrhythmia'
avdeftyp='Type of atrioventricular conduction defect'

```

```

axisab='Axis abnormality'
abhr='Abnormal heartrate'
longpr='Long P-R interval'
qwvab='Minnesota Q-wave code'
stwvab='ST or T wave abnormality'
qmi='Q-wave myocardial infarction'
longqrs='Long QRS duration'
longqt='Long QT duration'
stsegdep='ST segment depression'
stsegelv='ST segment elevation'
twvitms='T-wave items'
vdefcod='Ventricular conduction defect code';

format techflag techf. arythtyp arythtf. avdeftyp avdeftf. axisab axisabf.
abhr abhrf. longpr ynfmt. qwvab stwvab waveabf.
qmi longqrs longqt ynfmt. stsegdep stsegelv twvitms waveabf.
vdefcod vdefcdf.;

run;
proc freq data=ecg;
  tables tech_prb*techflag
         arythtyp*arrhyth
         avdeftyp*av_def
         axisab*axis
         abhr*hr
         longpr*pr
         qwvab*q_ant*q_post*q_lat*q_inf

stwvab*std_ant*std_lat*std_inf*stl_ant*stl_lat*stl_inf*t_ant*t_lat*t_inf
qmi*qwvab*stwvab
longqrs*qrs
longqt*qt
stsegdep*std_lat*std_inf*std_ant*qmi
stsegelv*stl_lat*stl_inf*stl_ant*qmi
twvitms*t_lat*t_inf*t_ant*qmi
vdefcod*vent_def / list missing;

run;
data calc.ecg(keep=habcid techflag arythtyp avdeftyp axisab abhr longpr qwvab
stwvab
                                qmi longqrs longqt stsegdep stsegelv twvitms vdefcod);

  merge ecg habc.ph;
  by habcid;
run;
proc contents data=calc.ecg;
title 'COntents of Y4 ECG calculated variables';
run;
proc freq data=calc.ecg;
  tables techflag arythtyp avdeftyp axisab abhr longpr qwvab stwvab
         qmi longqrs longqt stsegdep stsegelv twvitms vdefcod / missing;
run;

```

Appendix VII
Knee Pain Indicator Variables
Using Clinic Visit Workbook Variables

Investigator Name: Michael Nevitt

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4KP12MO	Knee pain (either knee) most days past 12 months	Knee pain in either knee lasting at least a month during last 12 months	Knee pain lasting at least a month during last 12 months (DAAJKMD=1)	Missing, Refused treated as No If (DAAJK12<0 or DAAJK12 in (7,8)) and DAAJKMD<0, then Y4KP12MO=.	0=No 1=Yes
Y4RKP12M	Right knee pain most days past 12 months	Right knee pain most days past 12 months	Knee pain lasting at least a month in right knee during last 12 months (DAAJKMD=1 and DAAJLRB1=1 or 3)	If DAAJKMD=1 and DAAJLRB1 is not in (1, 2, or 3), then Y4RKP12M=. If DAAJKMD≠1, then Y4RKP12M=0 If (DAAJK12<0 or DAAJK12 =7) and DAAJKMD<0, then Y4RKP12M=.	0=No 1=Yes
Y4LKP12M	Left knee pain most days past 12 months	Left knee pain most days past 12 months	Knee pain lasting at least a month in left knee during last 12 months (DAAJKMD=1 and DAAJLRB1=2 or 3)	If DAAJKMD=1 and DAAJLRB1 is not in (1, 2, or 3), then Y4LKP12M=. If DAAJKMD≠1, then Y4LKP12M=0 If (DAAJK12<0 or DAAJK12 =8) and DAAJKMD<0, then Y4LKP12M=.	0=No 1=Yes

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4RKP30D	Knee pain right knee most days in last 30	Knee pain in right knee most days in past 30	R knee pain most days in last 30 days (DAAJRKMS=1) If DAAJRKMS in (0,8) or DAAJRK30 in (0,8) or DAAJK12 in (0,8) then Y4RKP30D=0	If DAAJRKMS≠0 and DAAJRK30≠0 and DAAJK12≠0 and DAAJRKMS≠1, then Y4RKP30D=.	0=No 1=Yes
Y4LKP30D	Knee pain left knee most days in last 30	Knee pain in left knee most days in past 30	L knee pain most days in last 30 days (DAAJLKMS=1) If DAAJLKMS in (0,8) or DAAJLK30 in (0,8) or DAAJK12 in (0,8) then Y4LKP30D=0	If DAAJLKMS≠0 and DAAJLK30≠0 and DAAJK12≠0 and DAAJLKMS≠1, then Y4LKP30D=.	0=No 1=Yes
Y4KP30DA	Knee pain either knee most days in last 30	Knee pain (either) most days in past 30	If Y4RKP30D=1 or Y4LKP30D=1 then Y4KP30D=1; otherwise If Y4RKP30D=0 or Y4LKP30D=0 then Y4KP30DA=0	If Y4RKP30D≠0 and Y4LKP30D≠0 and Y4RKP30D≠1 and Y4LKP30D≠1, then Y4KP30DA=.	0=No 1=Yes
Y4RKPACT	Right knee pain with activity	Moderate or worse right knee pain in past 30 days on one or more activities in WOMAC scale	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in right knee during last 30 days when: walking on flat surface (DAAJRKFS), going up or down stairs (DAAJRKST), at night in bed (DAAJRKBD), standing upright (DAAJRKUP), getting in/out of chair (DAAJRKCH), getting in/out of car (DAAJRKIN)	If >2 of variables are missing and DAAJRK30 not in (0,8) and DAAJK12 not in (0,8), then Y4RKPACT is set to missing Missings and don't knows set to 0 for each component variable unless all of them are missing.	0=No 1=Yes

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4LKPACT	Left knee pain with activity	Moderate or worse left knee pain in past 30 days on one or more activities in WOMAC scale	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in left knee during last 30 days when: walking on flat surface (DAAJLKFS), going up or down stairs (DAAJLKST), at night in bed (DAAJLKBD), standing upright (DAAJLKUP), getting in/out of chair (DAAJLKCH), getting in/out of car (DAAJLKIN)	If >2 of variables are missing and DAAJLK30 not in (0,8) and DAAJK12 not in (0,8), then Y4LKPACT is set to missing Missings and don't knows set to 0 for each component variable unless all of them are missing.	0=No 1=Yes
Y4KPACT	Knee pain with activity (either knee)	Moderate or worse pain with activity in either knee	If Y4RKPACT=1 or Y4LKPACT=1 then Y4KPACT=1; otherwise If Y4RKPACT=0 or Y4LKPACT=0 then Y4KPACT=0	If Y4RKPACT≠0 and Y4LKPACT≠0 and Y4RKPACT≠1 and Y4LKPACT≠1, then Y4KPACT=.	0=No 1=Yes
Y4RKPSIG	Knee pain case - right knee	Right knee meets case definition	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (DAAJRKFS), going up or down stairs (DAAJRKST), at night in bed (DAAJRKBD), standing upright (DAAJRKUP), getting in/out of chair (DAAJRKCH), getting in/out of car (DAAJRKIN) or knee pain lasting ≥1month in past 12 months (DAAJMD=1) <u>and</u> DAAJLRB1=1 or 3, or R knee pain most days in past month (DAAJRKMS=1), then Y4RKPSIG=1. Else Y4RKPSIG=0	If DAAJRK12≤.z and DAAJRK30≤.z, then Y4RKPSIG=.	0=No 1=Yes

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4LKPSIG	Knee pain case - left knee	Left knee meets case definition	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (DAAJLKFS), going up or down stairs (DAAJLKST), at night in bed (DAAJLKBD), standing upright (DAAJLKUP), getting in/out of chair (DAAJLKCH), getting in/out of car (DAAJLKIN) or knee pain lasting ≥ 1 month in past 12 months (DAAJMD=1) <u>and</u> DAAJLRB1=2 or 3, or L knee pain most days in past month (DAAJLKMS=1), then Y4LKPSIG=1. Else Y4LKPSIG=0	If DAAJLK12 \leq .z and DAAJLK30 \leq .z, then Y4LKPSIG=.	0=No 1=Yes
Y4KPSIG	Knee pain case - either knee	Either knee meets case definition	If Y4RKPSIG=1 or Y4LKPSIG=1, then Y4KPSIG=1; If Y4RKPSIG=0 and Y4LKPSIG=0, then Y4KPSIG=0	If Y4RKPSIG=. and Y4LKPSIG=. then Y4KPSIG=.	0=No 1=Yes
Y4RWOMAC*	Right knee activity pain - Likert scale	Sum of right knee activity pain scores for 6 activities in WOMAC	sum of (DAAJRKFS, DAAJRKST, DAAJRKBD, DAAJKRUP, DAAJRKCH, DAAJRKIN)	If >2 of variables are missing, then Y4RWOMAC is set to Missing, otherwise Missing or Don't know set to 0 (No) before summing and then Y4RWOMAC is pro-rated	
Y4LWOMAC*	Left knee activity pain - Likert scale	Sum of left knee activity pain scores for 6 activities in WOMAC	sum of (DAAJLKFS, DAAJLKST, DAAJLKBD, DAAJKRUP, DAAJLKCH, DAAJLKIN)	If >2 of variables are missing, then Y4LWOMAC is set to Missing, otherwise Missing or Don't know set to 0 (No) before summing and then Y4LWOMAC is prorated	

* NOTE: Only the participants who have had knee pain in the past 12 months lasting at least one month go on to answer the activity pain items in Year 1. In later years, the criteria for who goes on to answer the activity pain items are loosened and a larger group of participants have a WOMAC score for each knee.

Documentation date: 2010-10-01

Knee Pain Indicator Variables
Using Core Home Visit Workbook (Version 1.2) Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4KP12MO	Knee pain (either knee) most days past 12 months	Knee pain in either knee lasting at least a month during last 12 months	Knee pain lasting at least a month during last 12 months (ZCAJKMD=1)	Missing, Don't Know, or Refused is treated as No If (ZCAJK12<0 or ZCAJK12=7) and ZCAJKMD<0, then Y4KP12MO=.	0=No 1=Yes
Y4KP30DA	Knee pain either knee most days in last 30	Knee pain (either) most days in past 30	Knee pain most days in last 30 days (ZCAJKMS=1) If ZCAJKMS in (0,8) or ZCAJK30 in (0,8) or ZCAJK12 in (0,8) then Y4KP30DA=0	otherwise Y4KP30DA=.	0=No 1=Yes
Y4LKP30D	Left knee pain knee most days in last 30	Left knee pain most days in past 30	N/A because the questions are asked differently and could not be calculated		
Y4RKP30D	Right knee pain knee most days in last 30	Right knee pain most days in past 30	N/A because the questions are asked differently and could not be calculated		
Y4KPACT	Knee pain with activity (either knee)	Moderate or worse pain with activity in either knee	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (ZCAJKFS), going up or down stairs (ZCAJKST), at night in bed (ZCAJKBD), standing upright (ZCAJKUP), getting in/out of chair (ZCAJKCH), getting in/out of car (ZCAJKIN)	If >2 of variables are missing and ZCAJK30 not in (0,8) and ZCAJK12 not in (0,8), then Y4KPACT is set to missing Otherwise, Y4KPACT=0	0=No 1=Yes

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4KPSIG	Knee pain case - either knee	Either knee meets case definition	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (ZCAJKFS), going up or down stairs (ZCAJKST), at night in bed (ZCAJKBD), standing upright (ZCAJKUP), getting in/out of chair (ZCAJKCH), getting in/out of car (ZCAJKIN) or knee pain lasting ≥ 1 month in past 12 months (ZCAJMD=1), or knee pain most days in past month (ZCAJKMS=1) <u>and</u> ZCAJLRB1=1,2 or 3 or ZCAJLRB2=1,2 or 3, then Y4KPSIG=1. Else Y4KPSIG=0	If ZCAJK12<=.z or ZCAJK12=7 then Y4KPSIG=.	0=No 1=Yes
Y4LKP12M	Left knee pain most days past 12 months	Left knee pain most days past 12 months	Knee pain lasting at least a month in left knee during last 12 months (ZCAJKMD=1 and ZCAJLRB1=2 or 3)	If ZCAJKMD=1 and ZCAJLRB1 is not in (1, 2, or 3), then Y4LKP12M=. If ZCAJKMD \neq 1, then Y4LKP12M=0 If (ZCAJK12<0 or ZCAJK12=7) and ZCAJKMD<0, then Y4LKP12M=.	0=No 1=Yes
Y4LKPACT	Left knee pain with activity	Moderate or worse left knee pain in past 30 days on one or more activities in WOMAC scale	Y4KPACT=1 and (ZCAJLRB2=2 or 3)	If Y4KPACT= 0, then Y4LKPACT=0. Otherwise Y4LKPACT is missing	0=No 1=Yes

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4LPLIK*	Left knee activity pain - Likert scale	Sum of left knee activity pain scores for 6 activities in WOMAC	sum of (ZCAJKFS, ZCAJKST, ZCAJKBD, ZCAJKUP, ZCAJKCH, ZCAJKIN) if ZCAJLRB2=2 or 3	If >2 of variables are missing, then Y4LPLIK is set to Missing, otherwise Missing, Don't know, Refused set to 0 (No) before summing and then Y4LPLIK is pro-rated If ZCAJK30 in (0,8) or ZCAJK12 in (0,8) then Y4LPLIK=0;	0-24
Y4LKPSIG	Left knee pain case	Indicator variable for knee pain case definition (left knee)	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (ZCAJKFS), going up or down stairs (ZCAJKST), at night in bed (ZCAJKBD), standing upright (ZCAJKUP), getting in/out of chair (ZCAJKCH), getting in/out of car (ZCAJKN) or knee pain lasting ≥ 1 month in past 12 months (ZCAJMD=1), or knee pain most days in past month (ZCAJKMS=1) <u>and</u> ZCAJLRB1=2 or 3 or ZCAJLRB2=2 or 3, then Y4LKPSIG=1. Else Y4LKPSIG=0	If ZCAJLRB1=8) <u>and</u> ZCAJLRB2=8), then Y4LKPSIG=. If ZCAJK12<=.z or ZCAJK12=7 then Y4LKPSIG=.	0=No 1=Yes

* Note: The Core Home Visit version of the activity questions (Q34 page 16 of the Revised Core Home Visit Workbook) is asked for both knees together, and then the participant is asked if the pain is in the right knee only, left knee only, or both left and right. If the participant responds “both left and right,” Y4LPLIK must equal Y4RPLIK. Otherwise, one can be 0 and the other non-zero. For this reason, this score is not completely comparable to the variables YxLWOMAC and YxRWOMAC and has been renamed to avoid confusion.

Documentation date: 2010-10-01

Page updated: 2009-06-25

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4RKP12M	Right knee pain most days past 12 months	Right knee pain most days past 12 months	Knee pain lasting at least a month in right knee during last 12 months (ZCAJKMD=1 and ZCAJLRB1=1 or 3)	If ZCAJKMD=1 and ZCAJLRB1 is not in (1, 2, or 3), then Y4RKP12M=. If ZCAJKMD≠1, then Y4RKP12M=0 If (ZCAJK12<0 or ZCAJK12 =7) and ZCAJKMD<0, then Y4RKP12M=.	0=No 1=Yes
Y4RKPACT	Right knee pain with activity	Moderate or worse right knee pain in past 30 days on one or more activities in WOMAC scale	Y4KPACT=1 and (ZCAJLRB2=1 or 3)	If Y4KPACT=0, then Y4RKPACT=0. Otherwise Y4RKPACT is missing	0=No 1=Yes
Y4RPLIK*	Right knee activity pain - Likert scale	Sum of right knee activity pain scores for 6 activities in WOMAC	sum of (ZCAJKFS, ZCAJKST, ZCAJKBD, ZCAJKUP, ZCAJKCH, ZCAJKIN) if ZCAJLRB2=1 or 3	If >2 of variables are missing, then Y4RPLIK is set to Missing; otherwise Missing, Don't know, Refused set to 0 (No) before summing and then Y4RPLIK is pro-rated If ZCAJK30 in (0,8) or ZCAJK12 in (0,8) then Y4RPLIK=0;	0-24

* Note: The Core Home Visit version of the activity questions (Q34 page 16 of the Revised Core Home Visit Workbook) is asked for both knees together, and then the participant is asked if the pain is in the right knee only, left knee only, or both left and right. If the participant responds “both left and right,” Y4LPLIK must equal Y4RPLIK. Otherwise, one can be 0 and the other non-zero. For this reason, this score is not completely comparable to the variables YxLWOMAC and YxRWOMAC and has been renamed to avoid confusion.

Documentation date: 2010-10-01

Page updated: 2009-06-25

Knee Pain Indicator Variables

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4RKPSIG	Right knee pain case	Indicator variable for knee pain case definition (right knee)	If any of the following are marked moderate, severe, or extreme (2, 3, or 4): pain in knees during last 30 days when: walking on flat surface (ZCAJKFS), going up or down stairs (ZCAJKST), at night in bed (ZCAJKBD), standing upright (ZCAJKUP), getting in/out of chair (ZCAJKCH), getting in/out of car (ZCAJKN) or knee pain lasting ≥ 1 month in past 12 months (ZCAJMD=1), or knee pain most days in past month (ZCAJKMS=1) <u>and</u> ZCAJLRB1=1 or 3 or ZCAJLRB2=1 or 3, then Y4RKPSIG=1. Else Y4RKPSIG=0	If ZCAJLRB1=8 <u>and</u> ZCAJLRB2=8, then Y4RKPSIG=. If ZCAJK12<=.z or ZCAJK12=7 then Y4RKPSIG=.	0=No 1=Yes

```

OPTIONS PAGESIZE=60 LINESIZE=132 nocenter formchar='|----+|----+=|-\<>*'
NOFMterr;
TITLE 'Project: Health ABC';
TITLE2 'Program: Kneepain_y4.sas';

data pain;
  set current.y4clnvis(keep=habcid daajkmd daajrkms daajrkfs daajrkst daajrkbd
daajrkup daajrkch daajrkin
                                daajlkms daajlkfs daajlkst daajlkbd daajlkup
daajlkch daajlkin
                                daajk12 daajrk30 daajlk30 daajlrbl );

  if (daajkmd<0 and daajrkms<0 and daajrkfs<0 and daajrkst<0 and daajrkbd<0 and
daajrkup<0 and daajrkch<0 and daajrkin<0 and daajlkms<0 and daajlkfs<0 and
daajlkst<0 and daajlkbd<0 and daajlkup<0 and daajlkch<0 and daajlkin<0 and
daajk12<0 and daajrk30<0 and daajlk30<0 and daajlrbl<0) then delete;

  *** Calculate pain variables ***;

  /*** SIGNIFICANT KNEE PAIN ***/
  if (daajkmd=1 and daajlrbl in (1,3)) or
daajrkms=1 or
  (daajrkfs in(2,3,4) or
daajrkst in(2,3,4) or
daajrkbd in(2,3,4) or
daajrkup in(2,3,4) or
daajrkch in(2,3,4) or
daajrkin in(2,3,4))
then Y4RKPSIG=1;

  if (daajkmd=1 and daajlrbl in (2,3)) or
daajlkms=1 or
  (daajlkfs in(2,3,4) or
daajlkst in(2,3,4) or
daajlkbd in(2,3,4) or
daajlkup in(2,3,4) or
daajlkch in(2,3,4) or
daajlkin in(2,3,4))
then Y4LKPSIG=1;

  if y4rkpsig ne 1 then y4rkpsig=0;
  if y4lkpsig ne 1 then y4lkpsig=0;

  if y4rkpsig=1 or y4lkpsig=1 then Y4KPSIG=1; else
  if y4rkpsig=0 or y4lkpsig=0 then y4kpsig=0;

  if (daajk12<=.z or daajk12=7) and (daajrk30<=.z or daajrk30=7) then
y4rkpsig=.;
  if (daajk12<=.z or daajk12=7) and (daajlk30<=.z or daajlk30=7) then
y4lkpsig=.;
  if y4rkpsig=. and y4lkpsig=. then y4kpsig=.;

  /*** 12-MONTH KNEE PAIN ***/
  if daajkmd=1 then do;
    Y4KP12MO=1;
    if daajlrbl=1 then do;

```

```

        Y4RKP12M=1; Y4LKP12M=0;
    end; else
    if daajlrbl=2 then do;
        y4rkp12m=0; y4lkp12m=1;
    end; else
    if daajlrbl=3 then do;
        y4rkp12m=1; y4lkp12m=1;
    end; else
    do;
        y4rkp12m=.; y4lkp12m=.;
    end;
end;
else do;
    y4kp12mo=0; y4rkp12m=0; y4lkp12m=0;
end;
if (daajk12<0 or daajk12=7) and (daajkmd<0) then do;
    y4rkp12m=.; y4lkp12m=.; y4kp12mo=.;
end;

IF y4KP12mo=0 AND y4RKP12m=1 AND y4LKP12m=1 THEN PUT / habcid= y4kp12mo=
y4rkp12m= y4lkp12m=
/ ' ' daajlrbl=;

    /*** 30-DAY KNEE PAIN ***/
    if daajrkms=1 then
Y4RKP30D=1; else
    if (daajrkms in (0,8) or daajrk30 in (0,8) or daajk12 in (0,8)) then
y4rkp30d=0; else

y4rkp30d=.;
    if daajlkms=1 then
Y4LKP30D=1; else
    if (daajlkms in (0,8) or daajlk30 in (0,8) or daajk12 in (0,8)) then
y4lkp30d=0; else

y4lkp30d=.;

    if y4rkp30d=1 or y4lkp30d=1 then Y4KP30DA=1; else
    if y4rkp30d=0 or y4lkp30d=0 then y4kp30da=0; else
    if y4rkp30d=. and y4lkp30d=. then y4kp30da=.;

    /*** PAIN WITH ACTIVITY ***/
    /*** Count # of missings in activity pain scores before setting to 0 ***/;
    nrmiss=nmiss(daajrkfs,daajrkst,daajrkbd,daajrkup,daajrkch,daajrkin);
    if daajrkfs in (2,3,4) or
        daajrkst in (2,3,4) or
        daajrkbd in (2,3,4) or
        daajrkup in (2,3,4) or
        daajrkch in (2,3,4) or
        daajrkin in (2,3,4)
        then Y4RKPACT=1; else y4rkpact=0;
    if (daajrk30 not in (0,8) and daajk12 not in (0,8) and nrmiss>2) then
y4rkpact=.;

    nlmiss=nmiss(daajlkfs,daajlkst,daajlkbd,daajlkup,daajlkch,daajlkin);

```

```

if daajlkfs in (2,3,4) or
  daajlkst in (2,3,4) or
  daajlkbd in (2,3,4) or
  daajlkup in (2,3,4) or
  daajlkch in (2,3,4) or
  daajlkin in (2,3,4)
  then Y4LKPACT=1; else y4lkpact=0;
if (daajlk30 not in (0,8) and daajk12 not in (0,8) and nlmiss>2) then
y4lkpact=. ;

if y4rkpact=1 or y4lkpact=1 then Y4KPACT=1; else
if y4rkpact=0 or y4lkpact=0 then y4kpact=0; else
if y4rkpact=. and y4lkpact=. then y4kpact=. ;

*** Change missing and don't knows to 0 per M.Nevitt ***;
** (Unless the whole knee pain data is missing) **;
if not (daajk12<=.z or daajk12=7) then do;
  if daajrkfs not in(0,1,2,3,4) then daajrkfs=0;
  if daajrkst not in(0,1,2,3,4) then daajrkst=0;
  if daajrkbd not in(0,1,2,3,4) then daajrkbd=0;
  if daajrkup not in(0,1,2,3,4) then daajrkup=0;
  if daajrkch not in(0,1,2,3,4) then daajrkch=0;
  if daajrkin not in(0,1,2,3,4) then daajrkin=0;
end;
if not (daajk12<=.z or daajk12=7) then do;
  if daajlkfs not in(0,1,2,3,4) then daajlkfs=0;
  if daajlkst not in(0,1,2,3,4) then daajlkst=0;
  if daajlkbd not in(0,1,2,3,4) then daajlkbd=0;
  if daajlkup not in(0,1,2,3,4) then daajlkup=0;
  if daajlkch not in(0,1,2,3,4) then daajlkch=0;
  if daajlkin not in(0,1,2,3,4) then daajlkin=0;
end;

*** Create summary scores for activities ***;
Y4RWOMAC=sum(daajrkfs,daajrkst,daajrkbd,daajrkup,daajrkch,daajrkin);
Y4LWOMAC=sum(daajlkfs,daajlkst,daajlkbd,daajlkup,daajlkch,daajlkin);

if daajrk30=0 or daajk12=0 then y4rwomac=0;
if daajlk30=0 or daajk12=0 then y4lwomac=0;

if (daajrk30 ne 0 and daajk12 ne 0 and nrmiss>2) then y4rwomac=. ;
if (daajlk30 ne 0 and daajk12 ne 0 and nlmiss>2) then y4lwomac=. ;

if nrmiss in (1,2) then y4rwomac=(y4rwomac/(6-nrmiss))*6;
if nlmiss in (1,2) then y4lwomac=(y4lwomac/(6-nlmiss))*6;
run;

* Calculate for home visits *;
data home;
  set current.y4corehv(keep=habcid /*zctype*/
    zcajk12 zcajkm d zcajlrbl zcajk30 zcajkms
    zcajkfs zcajkst zcajkb d zcajkup zcajkch zcajkin zcajlr b2);

*if zctype=4;

*** Calculate pain variables ***;

```

```

if zcajkmd=1 or
  (zcajkms=1 or
   (zcajkfs in(2,3,4) or
    zcajkst in(2,3,4) or
    zcajkbd in(2,3,4) or
    zcajkup in(2,3,4) or
    zcajkch in(2,3,4) or
    zcajkin in(2,3,4)))
  then y4kpsig=1;
else y4kpsig=0;

*Assign to r or l knee for 1st criteria*;
if zcajkmd=1 then do;
  if zcajlrbl=1 then do;
    lkpsig1=0;
    rkpsig1=1;
  end; else
  if zcajlrbl=2 then do;
    lkpsig1=1;
    rkpsig1=0;
  end; else
  if zcajlrbl=3 then do;
    lkpsig1=1;
    rkpsig1=1;
  end; else
  if zcajlrbl=8 then do;
    lkpsig1=.;
    rkpsig1=.;
  end;
end;

*Assign to r or l knee for 2nd criteria*;
if (zcajkms=1 or
  (zcajkfs in(2,3,4) or
   zcajkst in(2,3,4) or
   zcajkbd in(2,3,4) or
   zcajkup in(2,3,4) or
   zcajkch in(2,3,4) or
   zcajkin in(2,3,4))) then do;
  if zcajlrbl=1 then do;
    lkpsig2=0;
    rkpsig2=1;
  end; else
  if zcajlrbl=2 then do;
    lkpsig2=1;
    rkpsig2=0;
  end; else
  if zcajlrbl=3 then do;
    lkpsig2=1;
    rkpsig2=1;
  end; else
  if zcajlrbl=8 then do;
    lkpsig2=.;
    rkpsig2=.;
  end;
end;
y4lkpsig=max(lkpsig1,lkpsig2);

```

```

y4rkpsig=max(rkpsig1,rkpsig2);
if zcajk12<=.z or zcajk12=7 then do;
  y4kpsig=.;
  y4lkpsig=.;
  y4rkpsig=.;
end;
if y4kpsig=0 then do;
  y4lkpsig=0;
  y4rkpsig=0;
end;

/**** 12-MONTH KNEE PAIN ****/
if zcajkm=1 then do;
  y4kp12mo=1;
  if zcajlr=1 then do;
    y4rkp12m=1; y4lkp12m=0;
  end; else
  if zcajlr=2 then do;
    y4rkp12m=0; y4lkp12m=1;
  end; else
  if zcajlr=3 then do;
    y4rkp12m=1; y4lkp12m=1;
  end; else
  do;
    y4rkp12m=.; y4lkp12m=.;
  end;
end;
else do;
  y4kp12mo=0; y4rkp12m=0; y4lkp12m=0;
end;
if (zcajk12<0 or zcajk12=7) and (zcajkm<0) then do;
  y4rkp12m=.; y4lkp12m=.; y4kp12mo=.;
end;

/**** 30-DAY KNEE PAIN ****/
if zcajkms=1 then y4kp30da=1;
else
  if zcajkms in (0,8) or zcajk30 in (0,8) or zcajk12 in (0,8) then y4kp30da=0;
else
  y4kp30da=.;

/**** this part starred out because these variables cant be calculated (see
top)
if y4kp30da in (.,0) then do;
  y4lkp30d=y4kp30da; y4rkp30d=y4kp30da;
end; else
if y4kp30da=1 then do;
  if zcajlr=1 then do;
    y4lkp30d=0; y4rkp30d=1;
  end; else
  if zcajlr=2 then do;
    y4lkp30d=1; y4rkp30d=0;
  end; else
  if zcajlr=3 then do;
    y4lkp30d=1; y4rkp30d=1;
  end; else do;
    y4lkp30d=.; y4rkp30d=.;
  end;
end;

```

```

end;
***/

*** Count # of missings in activity pain scores before setting to 0 ***;
nmiss=nmiss(zcajkfs, zcajkst, zcajkbd, zcajkup, zcajkch, zcajkin);

if zcajkfs in (2,3,4) or
   zcajkst in (2,3,4) or
   zcajkbd in (2,3,4) or
   zcajkup in (2,3,4) or
   zcajkch in (2,3,4) or
   zcajkin in (2,3,4)
   then y4kpact=1; else y4kpact=0;
if (zcajk30 not in (0,8) and zcajk12 not in (0,8) and nmiss>2) then y4kpact=.;

if zcajlr2=1 then do;
   y4rkpact=y4kpact;
   y4lkpact=0;
end; else
if zcajlr2=2 then do;
   y4lkpact=y4kpact;
   y4rkpact=0;
end; else
if zcajlr2=3 then do;
   y4rkpact=y4kpact;
   y4lkpact=y4kpact;
end;
if y4kpact=0 then do;
   y4rkpact=0; y4lkpact=0;
end;

*** Change missing and don't knows to 0 per M.Nevitt ***;
** (Unless the whole knee pain data is missing) **;
if not (zcajk12<=.z or zcajk12=7) then do;
   if zcajkfs not in(0,1,2,3,4) then zcajkfs=0;
   if zcajkst not in(0,1,2,3,4) then zcajkst=0;
   if zcajkbd not in(0,1,2,3,4) then zcajkbd=0;
   if zcajkup not in(0,1,2,3,4) then zcajkup=0;
   if zcajkch not in(0,1,2,3,4) then zcajkch=0;
   if zcajkin not in(0,1,2,3,4) then zcajkin=0;
end;

*** Create summary scores for activities ***;
y4womac=sum(zcajkfs, zcajkst, zcajkbd, zcajkup, zcajkch, zcajkin);

if zcajk30 in (0,8) or zcajk12 in (0,8) then do;
   y4womac=0; y4rwomac=0; y4lwomac=0;
end; else
if (zcajk30 not in (0,8) and zcajk12 not in (0,8) and nmiss>2) then y4womac=.;
else
if nmiss in (1,2) then y4womac=(y4womac/(6-nmiss))*6;

if zcajlr2=1 then do;
   y4rwomac=y4womac; y4lwomac=0; end; else
if zcajlr2=2 then do;
   y4rwomac=0; y4lwomac=y4womac; end; else
if zcajlr2=3 then do;

```

```

        y4rwomac=y4womac; y4lwomac=y4womac; end; else
    if zcajlr2=8 then do;
        y4rwomac=0; y4lwomac=0; end;
    if y4lkpact=0 and y4lwomac=. then y4lwomac=0;
    if y4rkpact=0 and y4rwomac=. then y4rwomac=0;

    ** Home visit activity vars are not knee-specific, so are renamed **;
    rename y4lwomac=Y4LPLIK y4rwomac=Y4RPLIK;
run;

data kneepain(keep=habcid
              y4kpsig y4rkpsig y4lkpsig
              y4kp12mo y4rkp12m y4lkp12m
              y4kp30da y4rkp30d y4lkp30d
              y4kpact y4rkpact y4lkpact
              y4rwomac y4lwomac y4rplik y4lplik );
set pain(in=inp) home(in=inh);
by habcid;
  *Fix for one ppt who was in both clinic and home data*;
  if habcid=5097 and inh then delete;
y4lwomac=ceil(y4lwomac);
y4rwomac=ceil(y4rwomac);
y4lplik =ceil(y4lplik);
y4rplik =ceil(y4rplik);

label  y4kpsig = 'Knee pain case - either knee'
       y4rkpsig = 'Right knee pain case'
       y4lkpsig = 'Left knee pain case'
       y4kp12mo = 'Knee pain (either) most days past 12 mo'
       y4rkp12m = 'Right knee pain most days past 12 mo'
       y4lkp12m = 'Left knee pain most days past 12 mo'
       y4kp30da = 'Knee pain (either) most days in past 30'
       y4rkp30d = 'Right knee pain most days in past 30'
       y4lkp30d = 'Left knee pain most days in past 30'
       y4kpact  = 'Knee pain with activity - either knee'
       y4rkpact = 'Right knee pain with activity'
       y4lkpact = 'Left knee pain with activity'
       y4lwomac = 'L knee activity pain, Likert (clinic)'
       y4rwomac = 'R knee activity pain, Likert (clinic)'
       y4lplik  = 'L knee activity pain, Likert (home)'
       y4rplik  = 'R knee activity pain, Likert (home)';
format y4kpsig y4rkpsig y4lkpsig y4kp12mo y4rkp12m y4lkp12m y4kp30da y4lkp30d
y4rkp30d
       y4kpact y4rkpact y4lkpact yndk.;
run;
/*
data dups;
  set kneepain;
  by habcid;
  if not(first.habcid and last.habcid);
run;
proc print data=dups;
  title 'Duplicate in final data';
run;
*/
data calc.y4knpain;
  merge kneepain habc.ph(keep=habcid);

```

```
    by habcid;
run;

proc contents data=calc.y4knpain;
title 'Contents of Y4 knee pain calculated vars';
run;
proc freq data=calc.y4knpain;
  tables y4kpsig*y4rkpsig*y4lkpsig
         y4kp12mo*y4rkp12m*y4lkp12m
         y4kp30da*y4rkp30d*y4lkp30d
         y4kpact*y4rkpact*y4lkpact
         y4rwomac y4lwomac y4rplik y4lplik
  / list missing;
run;
```

Appendix VIII
Long Distance Corridor Walk Calculated Variables

Investigator Name: Eleanor Simonsick

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EXCLUDE4	LDCW exclusion code	Variable to capture who was excluded from all walks and why. Note: the first exclusion met is the one that is coded in EXCLUDE4. This year there were no participants excluded solely because no ECG was done, so EXCLUDE4=6 was omitted	If VTYPE GT 0 then EXCLUDE4=9; If VTYPE=0 then DO; EXCLUDE4=0; If EXCLUDE4=0 and D3MARQ=1 then EXCLUDE4=1; If EXCLUDE4=0 and D3HR40=1 then EXCLUDE4=5; If EXCLUDE4=0 and (D3SYSYN=1 or D3SYDIYN=1) then EXCLUDE4=2; If EXCLUDE4=0 and (D3WKAID2=1 or D120M=0) then EXCLUDE4=8; If EXCLUDE4=0 and (D3HA=1 or D3ANG=1 or D3HS=1) then EXCLUDE4=3; If EXCLUDE4=0 and (D3CP=1 or D3ANG2=1) then EXCLUDE4=4; If EXCLUDE4=0 and ((D1LD=0 or D1LD=2) and D32LAP<0) then EXCLUDE4=7;	A few participants who met exclusion criteria but were nonetheless tested were given a code 0 for EXCLUDE1 (affects 5 ppts)	0=not excluded 1=ECG abnormality 2=elevated standing bp 3=cardiac surgery 4=worsening cardiac symptoms 5=elevated heart rate, excluded 400m only 7=no data, reason unknown 8=mobility related exclusion 9=No clinic visit

Long Distance Corridor Walk Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
HOWFARY4 (analogous to WHENSTOP in Year 1)	LDCW completion status	Variable to capture who stopped when	If EXCLUDE4=1 or EXCLUDE4=2 or EXCLUDE4=3 or EXCLUDE4=5 or EXCLUDE4=6 or EXCLUDE4=7 or EXCLUDE4=8 or EXCLUDE4=9 then HOWFARY4=0; If EXCLUDE4=0 or EXCLUDE4=4 then DO: If D3C2MW=0 then HOWFARY4=2; If D3C2MW=1 and (D34LAP LE 0) then HOWFARY4=3; If D3C2MW=1 and (0<D34LAP<10) then HOWFARY4=4; If D34LAP=10 then HOWFARY4=5;		0=excluded 2=2-min walk not completed 3=2-min walk completed, 400M not attempted 4=400M walk not completed 5=completed 400M walk
D32SUM (analogous to P32SUMI in Year 1, but no imputation necessary)	2 minute walk distance	2 minute walk total distance (40xlaps + meters) whether completed or not	D32SUM=(40*D32LAP) + D32MTR; If D32LAP ≥ 0 AND D32MTR<0 THEN D32SUM=40*D32LAP		meters
D34SUM	400M walk distance	400-meter walk total distance (whether complete or not)	D34SUM=(40*D34LAP) + D34MTR; If D34LAP ≥ 0 AND D34MTR<0 THEN D34SUM=40*D34LAP		meters
Y4DID2MW (analogous to DID2MINW in Year 1)	completed 2 minute walk yes/no	Dichotomous variable for completing 2 min walk	If HOWFARY4 ≥ 3 then Y4DID2MW=1; ELSE Y4DID2MW=0	If HOWFARY4 is missing, then Y4DID2MW is missing	0=No 1=Yes

Long Distance Corridor Walk Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
Y4V2MINM (analogous to V2MINMTR in Year 1)	meters walked in 2 min - complete only	Number of meters walked for those who completed the 2 min walk only	If Y4DID2MW=1 then Y4V2MINM=D32SUM;	If Y4DID2MW=0 then Y4V2MINM=.	meters
Y4DID400 (analogous to DID400MW in Year 1)	completed 400m walk yes/no	Dichotomous variable for completing 400 m walk	If HOWFARY4=5 then DID400MW=1; ELSE DID400MW=0	If HOWFARY4<5, then DID400MW = 0	0=No 1=Yes
Y4V400MT (analogous to V400TIME in Year 1)	time to walk 400m - complete only	Time to walk 400 m for only those who completed the 400 meter walk	IF HOWFARY4=5 THEN Y4V400MT=D34TIME	If HOWFARY4<5, then Y4V400MT is missing	sec
Y4400MSD (analogous to MTR400SD in Year 1)	walking speed over 400m-completed only	Walking speed over 400 meters for only those who completed 400 meters	If WHENSTOP=5 then MTR400SD=ROUND(400/P34TIMEI, .01); ELSE MTR400SD=.	If WHENSTOP<5, then MTR400SD is missing, uses imputed version of distance	m/sec

```

/*This program creates the summary variables for the LDCW in Year 4. To the
extent possible all 3,075
participants will have a code for each status variable - EXCLUDE4, HOWFARY4,
Y4DID2MW, Y4DID400;
*/

```

```

OPTIONS PAGESIZE=50 LINESIZE=96 NODATE NONUMBER NOFMterr;
*create dataset similar to Eleanor's Y1Y4ldcw;
*added alternative code for source data (from "current" lib 9/24/02 EK);
data temp (keep=habcid vtype);
    set habc.visits;
    where visit=4;
*set current.y4clnvis (keep=habcid visitype);
*vtype=visitype;
run;
proc sort data=temp;
    by habcid;
run;

data temptemp;
    set habc4.d1(keep=habcid d120M d11d);
*set current.y4clnvis (keep=habcid d120M d11d);
run;

data d3;
    set habc4.d3 (keep=habcid d3marq d3hr1 d3af d3wpw d3ir d3vt d3av d3twave
d3hr2 d3hr40
                                d3sysb2 d3sysyn d3diab2 d3sydiyn d3wkaid2 d3ha d3ang
d3hs d3cp
                                d3angi d3hr2mw d3b2pl d3pls2 d32lap d32mtr
d32bpm d32pls d3c2mw
                                d3pex d3pcp d3psob d3pf d3pkp d3php d3pcf d3pbp
d34lap d3claps
                                d3addms d34time d3xcd d34bpm d34sys d34dia
d34hr d3cm4mw d34pex
                                d34phr d34pcp d34psob d34pf d34pkp d34php
d34pcf d34pbp d34prf
                                d34oth d3wcp d3wsob d3wkp d3whp d3wcf d3wfp
d3wnumb d3wlc d3wbp
                                d3woth);
*set current.y4clnvis (keep=habcid d3marq d3hr1 d3af d3wpw d3ir d3vt d3av
d3twave d3hr2 d3hr40
                                /*d3sysb2*/ d3sysyn /*d3diab2*/ d3sydiyn d3wkaid2 d3ha
d3ang d3hs d3cp
                                d3angi d3hr2mw d3b2pl d3pls2 d32lap d32mtr
d32bpm d32pls d3c2mw
                                d3pex d3pcp d3psob d3pf d3pkp d3php d3pcf d3pbp
d34lap d3claps
                                d3addms d34time d3xcd d34bpm d34sys d34dia
d34hr d3cm4mw d34pex
                                d34phr d34pcp d34psob d34pf d34pkp d34php
d34pcf d34pbp d34prf
                                d34oth d3wcp d3wsob d3wkp d3whp d3wcf d3wfp
d3wnumb d3wlc d3wbp
                                d3woth);
run;
data temp4;

```

```

merge temp d3 temptemp;
by habcid;
run;

data temptemp(keep=habcid d320stp1 d320tmls d320stp2 d320tm2s
              d320mw1 d320mw2);
set habc4.d3;
*set current.y4clnvis (keep=habcid d320stp1 d320tmls d320stp2 d320tm2s
                      d320mw1 d320mw2);
run;

data temp5;
merge temp4 temptemp;
by habcid;
run;

/*proc contents data=temp5;
run;*/

data calc.yly4cw (keep= habcid y4did2mw y4did400 exclude4 d32sum d34sum y4v2minm
y4v400mt howfary4 y4400msd);
set temp5;
LABEL EXCLUDE4 = 'LDCW exclusion code'
      D32SUM = '2 minute walk distance'
      D34SUM = '400m walk distance'
      HOWFARY4 = 'LDCW completion status'
      Y4DID2MW = 'completed 2 minute walk yes/no'
      Y4DID400 = 'completed 400m walk yes/no'
      Y4V2MINM = 'meters walked in 2 min - complete only'
      Y4V400MT = 'time to walk 400m - complete only'
      Y4400MSD = 'walking speed over 400m - completed only';
format Y4DID2MW Y4DID400 yndk. EXCLUD4 exclud4x. HOWFARY4 stop.;
*DATA JCR.Y1Y4CW; *SET JCR.Y1Y4LDCW;
*DETERMINATION OF LDCW EXCLUSION STATUS IN Y4 - EXCLUDE4. THIS IS ANALOGOUS TO
EXCLUDE2 USED
IN Y2 WITH A NEW TWIST - PERSONS WHO REPORTED INCREASED CARDIAC SYMPTOMS WERE NO
LONGER
EXCLUDED FROM ATTEMPTING THE 2-MINUTE WALK COMPONENT OF THE LDCW. IN THIS CASE
WE WILL RETAIN
CODE 4 TO INDICATE THE PRESENCE OF CARDIAC SYMPTOMS, BUT MODIFY THE LABEL.
0=NOT EXCLUDED
1=YES, Y4 ECG ABNORMALITY
2=YES, ELEVATED BP
3=YES, RECENT CARDIAC EVENT OR SURGERY
4=YES FROM 400M ONLY, INCREASED CARDIAC SYMPTOMS
5=YES, ELEVATED HR
7=REFUSED OR NO TEST, REASON UNKNOWN
8=YES,MOBILITY-RELATED - WALKING AID REQUIRED OR UNABLE TO COMPLETE 20M WALKS
9=NO CLINIC VISIT;
IF VTYPE not in (0,11) THEN EXCLUDE4=9;
IF VTYPE in (0,11) THEN DO;
EXCLUDE4=0;
IF EXCLUDE4=0 AND D3MARQ=1 THEN EXCLUDE4=1;
IF EXCLUDE4=0 AND D3HR40=1 THEN EXCLUDE4=5;
IF EXCLUDE4=0 AND (D3SYSYN=1 OR D3SYDIYN=1) THEN EXCLUDE4=2;
IF EXCLUDE4=0 AND (D3WKAID2=1 OR D120M=0) THEN EXCLUDE4=8;
IF EXCLUDE4=0 AND (D3HA=1 OR D3ANG=1 OR D3HS=1) THEN EXCLUDE4=3;

```

```

IF EXCLUDE4=0 AND ((D1LD=0 OR D1LD=2) AND D32LAP<0) THEN EXCLUDE4=7;
IF EXCLUDE4=0 AND (D3CP=1 OR D3ANGI=1) THEN EXCLUDE4=4;
END;
*IDS THAT MEET EXCLUSION CRITERIA BUT WERE TESTED 2247 5076 5140 5416 5505;
IF HABCID=2247 OR HABCID=5076 OR HABCID=5140 OR HABCID=5416 OR HABCID=5505 THEN
EXCLUDE4=0;
*deal with 2506 who did not do 2-min or LDCW, according to checklist for "other
reason";
IF HABCID=2506 THEN EXCLUDE4=7;
*deal with 2022 who refused LDCW and 2-min walk, according to checklist;
IF HABCID=2022 THEN EXCLUD4=7;
*DETERMINATION OF COMPLETION STATUS
0=EXCLUDED
2=2-MINUTE WALK NOT COMPLETED
3=2-MINUTE WALK COMPLETED, 400M NOT ATTEMPTED
4=400M NOT COMPLETED
5=COMPLETED 400M WALK;
D32SUM=(40*D32LAP) + D32MTR;
IF D32LAP GE 0 AND D32MTR<0 THEN D32SUM=40*D32LAP; *DISTANCE WALKED DURING 2
MINUTE WALK;

IF EXCLUDE4=1 OR EXCLUDE4=2 OR EXCLUDE4=3 OR EXCLUDE4=5
OR EXCLUDE4=6 OR EXCLUDE4=7 OR EXCLUDE4=8 OR EXCLUDE4=9 THEN HOWFARY4=0;
IF EXCLUDE4=0 OR EXCLUDE4=4 THEN DO;
IF D3C2MW=0 THEN HOWFARY4=2;
IF D3C2MW=1 AND (D34LAP LE 0) THEN HOWFARY4=3;
IF D3C2MW=1 AND (0<D34LAP<10) THEN HOWFARY4=4;
IF D34LAP=10 THEN HOWFARY4=5;
*deal with HB5434, who apparently finished 2-min walk, but d3c2mw is blank;
IF habcid=5434 and d3c2mw<0 then HOWFARY4=3;
END;

*Added D34SUM, for parallelism with year 1;
IF HOWFARY4>3 THEN DO;
D34SUM=(40*D34LAP) + D3ADDMS;
IF D34LAP GE 0 AND D3ADDMS<0 THEN D34SUM=40*D34LAP;end; *DISTANCE WALKED DURING
400m walk;

IF HOWFARY4 GE 3 THEN Y4DID2MW=1; ELSE Y4DID2MW=0;
IF Y4DID2MW=1 THEN Y4V2MINM=D32SUM;
IF HOWFARY4=5 THEN Y4DID400=1; ELSE Y4DID400=0;
IF HOWFARY4=5 THEN Y4V400MT=d34time;

*creating analog to MTR400SD in case Eleanor simply overlooked it;

if HOWFARY4=5 then Y4400MSD=ROUND(400/d34time,.01); else Y4400MSD=.;

*VARIABLE NAME COMPARISONS
      Y1          Y2          Y4
EXCLUDE1    EXCLUDE2    EXCLUDE4
WHENSTOP    HOWFARY2    HOWFARY4
DID2MINW    Y2DID2MW    Y4DID2MW
DID400MW    Y2DID400    Y4DID400
P32SUMI          B32SUM          D32SUM
N/A            B34SUM          D34SUM
V2MINMTR    Y2V2MINM    Y4V2MINM

```

```
V400TIME      Y2V400MT      Y4V400MT  
MTR400SD      Y2400MSD      Y4400MSD;
```

*All other summary variables derive from these;

```
run;  
/*  
proc contents data=calc.yly4cw;  
run;  
  
proc freq data=calc.yly4cw;  
    tables howfary4 exclude4 y4did2mw y4did400;  
run;  
  
proc sort data=calc.yly4cw;  
    by howfary4;  
run;  
  
proc univariate data=calc.yly4cw;  
var d34sum;  
run;  
  
proc print data=calc.yly4cw;  
    where howfary4<0;  
run;  
*/
```

Appendix IX
Performance Measure Calculated Variables

Investigator Name: Eleanor Simonsick

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
ABLE5CS	Did 5 chair stands	Indicator variable for able to do 5 chair stands. The scoring algorithm sets ABLE5CS to 0 if repeated stands were completed, but single chair stand coded as “rises using arms”	If single chair stand missing or refused and repeated stands are also missing, or if no suitable chair was available, or participant refused the single chair stand, then ABLE5CS is missing. If single chair is not attempted/unable, attempted and unable, or rises using arms, then ABLE5CS=0. If complete single stand, but not attempted/unable, refused, or unable to complete 5 stands, then ABLE5CS=0. If able to complete both single stand and 5 repeated stands, then ABLE5CS=1.		0=No 1=Yes
CHR5PACE	Chair stands per second	Chair stands per second	CHR5PACE=5 divided by the time to do 5 stands (in seconds) rounded to 1 decimal	If ABLE5CS=0 then CHR5PACE=0	sec ⁻¹
CAT5CS	EPESE score for chair stands		Round time for 5 stands to nearest decimal if ABLE5CS=0 then CAT5CS=0; if ABLE5CS=1 then: if time ≥16.7 then CAT5CS=1; if 13.7 ≤ time ≤ 16.6 then CAT5CS=2; if 11.2 ≤ time ≤ 13.6 then CAT5CS=3; if time ≤ 11.1 then CAT5CS=4	If ABLE5CS=1 and time <0 then CAT5CS=.	Scale from 0 (worst performance) to 4 (best performance)

Performance Measure Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
SBSCORE	EPESE score for standing balance	<p>EPESE categorical score for summary of standing balance results.</p> <p>The scoring algorithm reflects the EPESE administration procedure in which those who could not hold the semi-tandem stand for 10 seconds should not have moved on to the full tandem stand (even if they did have data for that test)</p>	<p>If refused or attempted/unable to do semi-tandem stand, or unable to hold for at least 1 second, then STSCORE=0.</p> <p>If holds for 1 to <10 seconds, then STSCORE=1.</p> <p>If at least 10 seconds, then STSCORE=2.</p> <p>Better of two tandem-stand trials is used.</p> <p>If refused or attempted/unable to do tandem stand, or unable to hold for at least 3 second then FTSCORE=0.</p> <p>If holds for 3 to <10 seconds, then FTSCORE=1.</p> <p>If at least 10 seconds then FTSCORE=2.</p> <p>If STSCORE in (0,1) then FTSCORE=0.</p> <p>SBSCORE=STSCORE + FTSCORE</p>	<p>If semi-tandem marked as holds for less than 30 seconds, but time is missing, assume held less than 10 seconds (STSCORE=1).</p> <p>If STSCORE is missing and tandem stand held for at least 10 seconds, then assume semi-tandem held for 30 (STSCORE=2)</p> <p>If semi-tandem is refused or unable then STSCORE=0</p> <p>If tandem is refused or unable then FTSCORE=0</p>	sec

Performance Measure Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
FSBTIME	Full standing balance test time 0-90	Total time of all stands. Those who could not hold the full tandem stand for 30 seconds were not supposed to be administered the one-leg stand so they were assigned 0 seconds for one-leg stand.	If STSCORE=0 then STTIME=0 If xxSTS*=3 then STTIME=30 If xxSTS*=2 then STTIME=xxSTSTM* If STSCORE=0 then FTTIME=0 If xxTS1*=3 or xxTS2=3, then FTTIME=30, otherwise FTTIME =max(xxTSTM , xxTS2TM). Otherwise if FTSCORE=0 then FTTIME=0. If semi-tandem scored as “Holds position between 1 and 29 seconds” and semi-tandem time is missing, use tandem time for semi-tandem time. If xxTR1=3 or xxTR2=3 then OLTIME=30. Otherwise, OLTIME=max(xxTR1TM*, xxTR2TM*). Otherwise OLTIME=0. If 0 < FTTIME < 30.00 then OLTIME=0. if STTIME=0 or FTTIME=0 then OLTIME=0. Round each balance time to two decimals FSBTIME=STTIME + FTTIME + OLTIME	If one-leg stand is refused or unable then OLTIME=0	sec
DID6MUW	Did 6-meter usual walk	Indicator variable for 6-meter usual walk (clinic visit only).	If steps or time is non-missing for the usual pace walk, then DID6MUW=1. Otherwise, if the walk was not attempted/ unable or attempted and unable to complete, then DID6MUW=0.	If home or proxy visit (no 6-meter walk component), DID6MUW=. If walk is missing or refused, then DID6MUW=.	0=No 1=Yes

*Clinic visit, core home visit, and proxy visit variables have different prefixes. Variables were all given standardized names first and then the calculated variables were created

Performance Measure Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
SIXMWTM	Time to walk 6m	The fastest of the two trials for the usual walk	SIXMWTM=minimum of non-missing usual walk times from two trials (for home and proxy home visits, a temporary variable, not kept, called Y4UWTM is created in the same way for the 3- or 4-meter walk)	If DID6MUW=0 or missing, SIXMWTM is missing	sec
Y4UWPACE	Walking speed (m/sec) over 3, 4 or 6m	Walking speed (m/sec) over 3,4 or 6m, using fastest trial	Divide SIXMWTM (or Y4UWTM) by number of meters walked (3, 4, or 6) and round to 2 decimals	If usual pace walk attempted/unable or attempted and unable to complete then Y4UWPACE=0	m/sec
Y4UWSCR	EPESE categories for walking speed	EPESE categories for usual walking speed, using fastest trial	If SIXMWTM (or Y4UWTM)=0 then Y4UWSCR=0; if $0 < Y4UWPACE \leq .46$ then Y4UWSCR=1; if $.47 \leq Y4UWPACE \leq .64$ then Y4UWSCR=2; if $.65 \leq Y4UWPACE \leq .82$ then Y4UWSCR=3; if $Y4UWPACE \geq .83$ then Y4UWSCR=4		Scale from 0 (worst performance) to 4 (best performance)
EPESEPPB	EPESE performance battery score 0-12	EPESE performance battery score 0-12**	EPESEPPB=Y4UWSCR+ CAT5CS + SBSCORE	If any of the component scales is missing, then EPESEPPB is missing	Scale from 0 (worst) to 12 (best performance)
DID6MNW	Did 6m narrow walk	Indicator variable for 6m narrow walk.	If DID6MUW=. then DID6MNW=. if DID6MUW=0 then DID6MNW=0; If attempted / unable to do the narrow walk then DID6MNW=0 If all attempted trials strayed outside the lines, then DID6MNW=0. if any of three trials stayed within bounds, then DID6MNW=1.	If a participant did the usual 6 meter walk, but did not attempt or refused the narrow (7), they were coded as unable (0)	0=No 1=Yes

**The scoring algorithm reflects the EPESE administration procedure in which those who could not hold the semi-tandem stand for 10 seconds did not move on to the full tandem stand. This is not how the examiners were instructed to administer the exam, however, so a number of participants had full-tandem stand data who did not hold the semi-tandem stand for 10 seconds

Performance Measure Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
NWTIME	Time to walk a 20cm wide 6m course	Time to walk a 20cm wide 6m course. Uses the fastest of three trials	NWTIME = minimum time from acceptable trial(s)	If DID6MNW=0 or missing, NWTIME is missing	Sec
NWPACE	Walking speed for narrow 6m walk	Walking speed (m/sec) over 6m narrow walk, using fastest trial	NWPACE = 6/NWTIME rounded to 2 decimals	If DID6MNW=0 then NWPACE=0	m/sec
NWSCORE	EPESE category for narrow walk	EPESE categories for narrow walk, using fastest trial	If DID6MNW=0 then NWSCORE=0; if $0 < \text{NWPACE} \leq .46$ then NWSCORE=1; if $.47 \leq \text{NWPACE} \leq .64$ then NWSCORE=2; if $.65 \leq \text{NWPACE} \leq .82$ then NWSCORE=3; if $\text{NWPACE} \geq .83$ then NWSCORE=4		Scale from 0 (worst performance) to 4 (best performance)
NWSCOREQ	% difference between narrow pace & usual walk pace	Categories based on percent difference between Y4UWPACE and NWPACE	PACEDIFF=Y4UWPACE-NWPACE; PCPACEDF=round(100*PACEDIFF/Y4UWPACE, .01); if PCPACEDF > 15 then NWSCOREQ=1; if $15 \geq \text{PCPACEDF} > 5$ then NWSCOREQ=2; if $5 \geq \text{PCPACEDF} \geq -5$ then NWSCOREQ=3; if $.z < \text{PCPACEDF} < -5$ then NWSCOREQ=4;	If DID6MNW=. then NWSCOREQ=. if DID6MNW=0 then NWSCOREQ=. if DID6MUW=0 and DID6MNW=0 then NWSCORE=.	%

Performance Measure Calculated Variables

Note all performances except for the standing balance measures were converted to pace (meters/sec or chair stands/sec) to allow those unable to do a test to be coded 0. It was difficult to determine whether the 7:Not attempted/refused were not administered a test or if they refused or did not attempt a test because of a physical problem.

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
CSRATIO	Chair stands performance ratio using a modification of the MacArthur approach***	A summary scale of chair stand performance. 1 chair stand per second was used for the denominator	$CSRATIO = \text{round}(CHR5PACE/1, .01)$; if $CSRATIO > 1$ then $CSRATIO = 1$	Chair stand performances was converted to pace - chair stands/sec - to allow those unable to do a test to be coded 0;	unitless
FSBRATIO	Standing balance performance ratio using a modification of the MacArthur approach***)	A summary scale of standing balance performance. 90 seconds (total time for all three stands) was used for the denominator	$FSBRATIO = FSBTIME/90$ rounded to 2 decimals	If FSBTIME is missing, then FSBRATIO is missing	unitless
Y4UWRATIO	Usual walk performance ratio using a modification of the MacArthur approach ***	A summary scale of usual walk performance. 2 m/sec was used for the denominator	$Y4UWRATIO = \text{round}(Y4UWPACE/2, .01)$ if $Y4UWRATIO > 1$ then $Y4UWRATIO = 1$	Usual walk performances was converted to pace - meters/sec - to allow those unable to do a test to be coded 0;	unitless
NWRATIO	Narrow walk performance ratio using a modification of the MacArthur approach ***	A summary scale of 6 meter narrow walk performance. 2 m/sec was used for the denominator	$NWRATIO = \text{round}(NWPACE/2, .01)$; if $NWRATIO > 1$ then $NWRATIO = 1$	Narrow walk performances was converted to pace - meters/sec - to allow those unable to do a test to be coded 0;	unitless

***Seeman et al. J Gerontol Med Sci 1994(49)M97-M108). With the MacArthur data, the 99th percentile was used as the denominator and scores above the 99th percentile were set to equal the 99th percentile. Here, the maximum attainable performance within an older cohort as used as the denominator.

Performance Measure Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
HABCPPB	Health ABC performance score 0-4	Summary measure combining usual walk, narrow walk, chair stand, and standing balance performance ratios	HABCPPB= FSBRATIO+CSRATIO+Y4UWRATIO+ NWRATIO	If any of the component scales is missing, then HABCPPB is missing	Scale from 0 (worst) to 4 (best performance)
PPB12CAT ***NOT FOR USE OTHER THAN FOR COMPARISON TO EPESEPPB (use EPESEPPB or HABCPPB)	Categorical scoring of HABCPPB 0-12	Categories of HABCPPB (summary measure combining usual walk, narrow walk, chair stand, and standing balance performance ratios)	If HABCPPB > 3.66 then PPB12CAT=12; If 3.33 < HABCPPB < 3.66 then PPB12CAT=11; If 3.0 < HABCPPB ≤ 3.33 then PPB12CAT=10; If 2.66 < HABCPPB ≤ 3.0 then PPB12CAT=9; If 2.33 < HABCPPB ≤ 2.66 then PPB12CAT=8; If 2.0 < HABCPPB ≤ 2.33 then PPB12CAT=7; If 1.66 < HABCPPB ≤ 2.0 then PPB12CAT=6; If 1.33 < HABCPPB ≤ 1.66 then PPB12CAT=5; If 1.0 < HABCPPB ≤ 1.33 then PPB12CAT=4; If 0.66 < HABCPPB ≤ 1.0 then PPB12CAT=3; If 0.33 < HABCPPB ≤ 0.66 then PPB12CAT=2; If 0.0 < HABCPPB ≤ 0.33 then PPB12CAT=1; If HABCPPB = 0 then PPB12CAT=0;	If HABCPPB is missing, then PPB12CAT is missing	Scale from 0 (worst performance) to 12 (best performance)

```

/*****
*****
* Program Name: Perform_Y4.sas
* Saved as:      \\Fu-hsing-c\habc\habc_sas\Calculated Variables\Programs\Year
4\Perform_Y4.sas
* Study:        HABC
* Purpose:      Creates calculated variables for Year 4 for
*              Standing Balance, Balance Walks, and Chair Stand
* Input:        y4ClnVis, Y4CoreHV, and Y4Proxy
* Output:       Perfm_Y4
* Programmer:   Ase Sewall (in coordination with Emily Kenyon)
* Date:         July 10, 2009
* Comments:     Consolidate and reconciled differences in various sets of code
*              (Consultation from Eleanor Simonsick)
*****
*****/;

*Set macro variable to use perfm_macros.sas;
%let Y=Y4;
%let pre=d3;
%let hpre=z4;
*****----- Combine clinic, home visit, and proxy data and rename variables
to single set of names -----*****;

data y4clnvis (keep=habcid d3scs d3rcs d3sec
                d3sts d3ststm d3ts2 d3ts1 d3ts2tm d3tstm
                d3trl d3tr2 d3trltm d3tr2tm
                d3uptm1 d3up1 d3upru1 d3uptm2 d3up2 d3upru2
                d320cna d320trl d320ctl d320tr2 d320ct2 d320tr3 d320ct3);
    set current.y4clnvis (where=(visitype in (0,10,11)));
run;

data y4corehv;
    set current.y4corehv (keep=habcid z4scs z4rcs z4sec
                            z4sts z4ststm z4ts2 z4ts1 z4ts2tm z4tstm
                            z4trl z4tr2 z4trltm z4tr2tm
                            z44mwtm1 z44mw1 z44mwtm2 z44mw
                            rename=(z4ts2tm=z4ts2tmX));
    *****Drop HB5097 (he has CV and CHV data and we want to only use the clinic
visit data);
    if habcid=5097 then delete;
    ** convert z4ts2tm to a numeric;
    length z4ts2tm 8;
    z4ts2tm = z4ts2tmX;
run;

data y4proxy;
    set current.y4proxy (keep=habcid ycscs yrcs ycsec
                            ycsts ycststm ycts2 ycts1 ycts2tm yctstm
                            yctrl yctr2 yctrltm yctr2tm
                            yc4mwtm1 yc4mw1 yc4mwtm2 yc4mw);
run;

data visit;
    merge y4proxy (in=c rename=(ycscs=scs yrcs=rcs
                                ycsts=sts ycststm=ststm ycts1=ts1 yctstm=tstm ycts2=ts2 ycts2tm=ts2tm
                                yctrl=trl yctrltm=trltm yctr2=tr2 yctr2tm=tr2tm ycsec=sec));

```

```
y4clnvis (in=a rename=(d3scs=scs d3rcs=rscs
    d3sts=sts d3ststm=ststm d3ts1=ts1 d3tstm=tstm d3ts2=ts2 d3ts2tm=ts2tm
    d3tr1=tr1 d3tr1tm=tr1tm d3tr2=tr2 d3tr2tm=tr2tm d3sec=sec))
y4corehv (in=b rename=(z4scs=scs z4rcs=rscs
    z4sts=sts z4ststm=ststm z4ts1=ts1 z4tstm=tstm z4ts2=ts2 z4ts2tm=ts2tm
    z4tr1=tr1 z4tr1tm=tr1tm z4tr2=tr2 z4tr2tm=tr2tm z4sec=sec));
by habcid;
if a then source="V";
else if b then source="H";
else if c then source="P";
run;

%include "&_prospath\Calculated Variables\Programs\perfm_macros.sas";
```

```

%macro perfm;
data Perf_&Y;
  set visit;

  ***** Standing Balance *****;

  *--- (1) Semi-tandem score;
  if STS in (7,9) then STSCORE=0; *Refused attempted/unable;
  else if STS<0 and STSTM<0 then STSCORE=.;
  else if (STS<0 and 0<STSTM < 1) or STS=1 then STSCORE=0; *Holds position
for less than one second;
  else if (STS=2 and STSTM<0 and (TS1=3 or TS2=3 or TSTM>0 or TS2TM>0))
then do;*Assumes ST time at least as good as FT;
  if TS1=3 or TS2=3 then STSTM=30;
  else STSTM=max(TSTM,TS2TM);
  end;
  if (STS=2 or STS<0) and 0<STSTM<10 then STSCORE=1; *Holds position for less
than 10 seconds;
  else if STS=2 and STSTM<0 then STSCORE=1; *If less than 30 sec and time
missing, assume less than 10 sec;
  else if STS=3 or ((STS=2 or STS<0)and STSTM ge 10) then STSCORE=2; *Holds
position for at least 10 seconds;

  *--- (2) Full-tandem score (1st attempt);
  if TS1 in (7,9) then FT1=0; *Refused attempted/unable;
  else if TS1<0 and TSTM<0 then FT1=.;
  else if (TS1<0 and 0<TSTM<1) or TS1=1 then FT1=0; *Holds position for less
than one second;
  else if (TS1=2 or TS1<0) and TSTM<10 then FT1=1; *Holds position for less
than 10 seconds (assumes missing time means less than 10 sec);
  else if TS1=3 or ((TS1=2 or TS1<0) and TSTM ge 10) then FT1=2; *Holds
position for at least 10 seconds;
  if (ts1=1 or (ts1=2 and 0 le tstm lt 3)) then FT1=0;*Full-tandem < 3
seconds;
  *--- (3) Full-tandem score (2nd attempt);
  if TS2 in (7,9)then FT2=0; *Refused attempted/unable;
  else if TS2<0 and TS2TM<0 then FT2=.;
  else if (TS2<0 and 0<TS2TM<1) or TS2=1 then FT2=0; *Holds position for
less than one second;
  else if (TS2=2 or TS2<0) and TS2TM<10 then FT2=1; *Holds position for less
than 10 seconds (assumes missing time means less than 10 sec);
  else if TS2=3 or ((TS2=2 or TS2<0) and TS2TM ge 10) then FT2=2; *Holds
position for at least 10 seconds;
  if (ts2=1 or (ts2=2 and 0 le ts2tm lt 3)) then FT2=0;
  FTSCORE=max(FT1,FT2); *Full-tandem score is better of two tries;

  *--- (4) Combined standing balance scores;
  if STSCORE in (0,1) then FTSCORE=0; *Should not go on to full-tandem if
unable to hold semi-tandem for 10 seconds;
  if STSCORE=. and FTSCORE>0 then STSCORE=2 ;*If semi-tandem test missing and
full-tandem exists, then assume completed at least 10 seconds on semi-tandem

  *--- (5) Final standing balance score;
  label SBSCORE ="EPESE SCORE FOR STANDING BALANCE";
  SBSCORE=STSCORE+FTSCORE;

```

```

*--- (6) For debug print only;
if 0 lt ststm lt 10 then STimelt10="<10";

***** Chair stands *****;

label ABLE5CS  ="DID 5 CHAIR STANDS YES=1";
IF (SCS=.M AND RCS=.A) OR (SCS=7 AND (RCS=.A OR RCS=.E)) THEN ABLE5CS=.;
  else IF SCS=1 THEN ABLE5CS=0;
  else IF SCS=2 AND (RCS=1 OR RCS=7 or RCS=9) THEN ABLE5CS=0;
  else IF SCS=9 THEN ABLE5CS=0;
  else IF SCS=3 THEN ABLE5CS=.;
  else IF SCS=0 THEN ABLE5CS=0;
  else IF SCS NE 1 AND RCS=2 THEN ABLE5CS=1;
  else IF SCS=7 THEN ABLE5CS=.;
  else IF SCS=2 AND RCS=2 THEN ABLE5CS=1;

label CAT5CS  ="EPESE SCORE FOR CHAIR STANDS";
RSEC=ROUND(SEC, .1);
IF ABLE5CS=0 THEN CAT5CS=0;
else IF ABLE5CS=1 then do;
  IF RSEC GE 16.7 THEN CAT5CS=1;
  IF 13.7 LE RSEC LE 16.6 THEN CAT5CS=2;
  IF 11.2 LE RSEC LE 13.6 THEN CAT5CS=3;
  IF 0<RSEC LE 11.1 THEN CAT5CS=4;
  else if RSEC lt 0 then CAT5CS=.;
end;

label CHR5PACE  ="CHAIR STANDS PER SECOND";
CHR5PACE=ROUND(5/SEC, .01);
IF ABLE5CS=0 THEN CHR5PACE=0;

label CSRATIO  ="CHAIR STANDS PERFORMANCE RATIO";
CSRATIO=ROUND(CHR5PACE/1, .01); IF CSRATIO GT 1 THEN CSRATIO=1;

**Total seconds to examine full standing balance assessment and create a
score
  from 0 to 90 - total time of all stands;

*--- (1) semi-tandem time;
IF STSCORE=0 THEN STTIME=0;
else do;
  IF STS=3 THEN STTIME=30;
  IF STS=2 THEN STTIME=STSTM;
  IF (STS=2 AND STSTM=.M) AND TS1=2 THEN STTIME=TSTM;
  IF (STS=2 AND STSTM=.M) AND (TS1=3 OR TS2=3) THEN STTIME=30;
end;
STTIME=ROUND(STTIME, .01);

*--- (2) full-tandem time;
if STSCORE in (0,1) then FTTIME=0;
  else IF TS1=3 OR TS2=3 THEN FTTIME=30;
  else IF TS1=2 or TS2=2 THEN FTTIME=MAX(TSTM, TS2TM);
  else if FTSCORE=0 then FTTIME=0;
*New code to try to be more consistent;
  if FTSCORE>0 and STTIME<0 then STTIME=FTTIME;

```

```

*--- (3) one-leg stand;
if tr1=3 or tr2=3 then OLTIME=30;
  else OLTIME=MAX(Tr1TM, Tr2TM);
  if FTTIME<0 and OLTIME>0 then FTTIME=OLTIME;
  if OLTIME<0 and (tr1 in (7,9) or (tr1=1 and tr2 in (7,9,1,.M))) then
OLTIME=0;*new;
if STTIME=0 or FTTIME=0 then OLTIME=0;
if 0 lt FTTIME lt 30 then OLTIME=0;

*--- (4) combined time for all three stands;
label FSBTIME ="Standing Balance Test Time (0-90)"
  FSBRATIO ="Standing Balance Time Ratio";
FSBTIME=STTIME + FTTIME + OLTIME;
FSBRATIO=round(FSBTIME/90, .01);

%if (&Y=Y1 or &Y=Y4 or &Y=Y6 or &Y=Y10 or &Y=Y11) %then %do;
***** Balance walks *****;

**Balance walk (6 Meter and Narrow);
/* Code to compute EPESE categorical score for the measured walk. It
is difficult to determine whether the 7s were not administered a test
of if they refused/did not attempt a test b/c a physical problem.
The INYN variable used in Y1 is not avail in other years*/;

*--- (Usual walk);

***DID6MUW (Did usual walk);
LABEL DID6MUW ="DID 6M USUAL WALK YES=1";
if Source="V" then do;
  IF &pre.UP1>0 or &pre.uptm1 gt .Z then DID&Y.UW=1;
  else IF &pre.UPRU1<=.z THEN DID&Y.UW=.;
  else if &pre.UPRU1=9 THEN DID&Y.UW=0;
  else IF &pre.UPRU1=7 THEN DID&Y.UW=.;
  DID6MUW=DID&Y.UW;
end;
%if (&Y ne Y1) %then %do;
if Source="H" then do;
  if (&hpre.4mwtm1 ge 0 and &hpre.4mwtm2 ge 0) then DID&Y.UW=1;
  else if &hpre.4mwl=9 or &hpre.4mwl=1 then DID&Y.UW=0;
end;
%end;
  %if (&Y=Y2 or &Y=Y3 or &Y=Y4 or &Y=Y5 or &Y=Y6 or &Y=Y8) %then %do;
  if Source="P" then do;
  if (yc4mwtm1 ge 0 and yc4mwtm2 ge 0) then DID&Y.UW=1;
  else if yc4mwl=9 or yc4mwl=1 then DID&Y.UW=0;
end;
%end;

***SIXMWTM (Time for 6m usual walk);

if Source="V" then do;
  IF &pre.UPTM1 > 0 AND &pre.UPTM2 > 0 THEN &Y.UWTM=MIN(&pre.UPTM1,
&pre.UPTM2);
  else if &pre.uptm1 lt 0 and &pre.uptm2 ge 0 then &Y.UWTM=&pre.uptm2;
  else if &pre.uptm1 ge 0 and &pre.uptm2 lt 0 then &Y.UWTM=&pre.uptm1;
  SIXMWTM=&Y.UWTM;
label SIXMWTM ="TIME TO WALK 6M";
end;

```

```

    %if (&Y ne Y1) %then %do;
if Source="H" then do;
    if (&hpre.4mwtm1 ge 0 and &hpre.4mwtm2 ge 0) then
&Y.UWTM=min(&hpre.4mwtm1, &hpre.4mwtm2);
    else if &hpre.4mwtm1 lt 0 and &hpre.4mwtm2 ge 0 then &Y.UWTM=&hpre.4mwtm2;
    else if &hpre.4mwtm1 ge 0 and &hpre.4mwtm2 lt 0 then &Y.UWTM=&hpre.4mwtm1;
end;
    %end;
    %if (&Y=Y2 or &Y=Y3 or &Y=Y4 or &Y=Y5 or &Y=Y6 or &Y=Y8) %then %do;
else if Source="P" then do;
    if (yc4mwtm1 ge 0 and yc4mwtm2 ge 0) then &Y.UWTM=min(yc4mwtm1, yc4mwtm2);
    else if yc4mwtm1 lt 0 and yc4mwtm2 ge 0 then &Y.UWTM=yc4mwtm2;
    else if yc4mwtm1 ge 0 and yc4mwtm2 lt 0 then &Y.UWTM=yc4mwtm1;
end;
    %end;
***&Y.UWPACE (pace for usual walk);
label &Y.UWPACE ="WALKING SPEED (M/SEC) OVER 3,4, or 6M";
    if Source="V" then do;
&Y.UWPACE=ROUND(6/&Y.UWTM, .01);
    IF DID&Y.UW=0 THEN &Y.UWPACE=0;
    end;
    %if (&Y ne Y1) %then %do;
if Source="H" then do;
    if &hpre.4mw=1 then &Y.UWPACE=round(4/&Y.UWTM, .01);
    if &hpre.4mw=2 then &Y.UWPACE=round(3/&Y.UWTM, .01);
end;
    %end;
    %if (&Y=Y2 or &Y=Y3 or &Y=Y4 or &Y=Y5 or &Y=Y6 or &Y=Y8) %then %do;
else if Source="P" then do;
    if yc4mw=1 then &Y.UWPACE=round(4/&Y.UWTM, .01);
    if yc4mw=2 then &Y.UWPACE=round(3/&Y.UWTM, .01);
end;
    %end;
***&Y.UWSCR (Score for usual walk);
label &Y.UWSCR ="EPESE SCORE FOR WALKING SPEED";
IF DID&Y.UW=0 THEN &Y.UWSCR=0;
    else IF 0 LT &Y.UWPACE LE .46 THEN &Y.UWSCR=1;
    else IF .47 LE &Y.UWPACE LE .64 THEN &Y.UWSCR=2;
    else IF .65 LE &Y.UWPACE LE .82 THEN &Y.UWSCR=3;
    else IF &Y.UWPACE GE .83 THEN &Y.UWSCR=4;

***&Y.UWRATIO (Ratio usual walk);
label &Y.UWRATIO ="USUAL WALK PERFORMANCE RATIO";
&Y.UWRATIO=ROUND(&Y.UWPACE/2, .01);*2 m/sec is considered fastest attainable
pace in this age group;
IF &Y.UWRATIO GT 1 THEN &Y.UWRATIO=1;

*--- (20cm narrow walk);

***DID6MNW (Did 20cm narrow walk);
label DID6MNW ="DID 6M NARROW WALK YES=1";
IF DID&Y.UW=. THEN DID6MNW=.;
else if DID6MUW=0 THEN DID6MNW=0;

else IF &pre.20TR1=0 AND &pre.20TR2=0 AND &pre.20TR3 < 1 THEN DID6MNW=0;
    IF &pre.20CNA in (0,9) THEN DID6MNW=0;*One version of the form coded not
attempted/unable as 0;

```

```

else IF DID6MUW=1 AND &pre.20CNA=7 THEN DID6MNW=0;
else IF &pre.20TR1=1 OR &pre.20TR2=1 OR &pre.20TR3=1 THEN DID6MNW=1;

***NWTIME***;
label NWTIME ="TIME TO WALK A 20CM WIDE 6M COURSE";
IF (&pre.20TR1=1 AND &pre.20CT1>0) THEN VT1T=1;
ELSE VT1T=0;
IF (&pre.20TR2=1 AND &pre.20CT2>0) THEN VT2T=1;
ELSE VT2T=0;
IF &pre.20TR3=1 AND &pre.20CT3>0 THEN VT3T=1;
ELSE VT3T=0;
IF DID6MNW=1 THEN DO;
IF VT1T=1 AND VT2T=0 AND VT3T=0 THEN NWTIME=&pre.20CT1;
else IF VT1T=0 AND VT2T=1 AND VT3T=0 THEN NWTIME=&pre.20CT2;
else IF VT1T=0 AND VT2T=0 AND VT3T=1 THEN NWTIME=&pre.20CT3;
else IF VT1T=1 AND VT2T=1 AND VT3T=0 THEN NWTIME=MIN(&pre.20CT1,
&pre.20CT2);
else IF VT1T=1 AND VT2T=0 AND VT3T=1 THEN NWTIME=MIN(&pre.20CT1,
&pre.20CT3);
else IF VT1T=0 AND VT2T=1 AND VT3T=1 THEN NWTIME=MIN(&pre.20CT2,
&pre.20CT3);
else IF VT1T=1 AND VT2T=1 AND VT3T=1 THEN NWTIME=MIN(&pre.20CT1,
&pre.20CT2, &pre.20CT3);
END;

***NWPACE***;
label NWPACE="WALKING SPEED FOR NARROW WALK 6M";
NWPACE=ROUND(6/NWTIME, .01);
IF DID6MNW=0 THEN NWPACE=0;

***NWSCORE***;
label NWSCORE ="EPESE CATEGORY FOR NARROW WALK";
IF DID6MNW=0 THEN NWSCORE=0;
else IF 0 LT NWPACE LE .46 THEN NWSCORE=1;
else IF .47 LE NWPACE LE .64 THEN NWSCORE=2;
else IF .65 LE NWPACE LE .82 THEN NWSCORE=3;
else IF NWPACE GE .83 THEN NWSCORE=4;

***NWRATIO***;
label NWRATIO ="NARROW WALK PERFORMANCE RATIO";
NWRATIO=ROUND(NWPACE/2, .01);
IF NWRATIO GT 1 THEN NWRATIO=1;

*--- (Comparison between 6M Usual walk 20cm Narrow walk);

if &Y.UWPACE gt 0 and nwpace gt 0 then do; /* Added by Ase,
changed to ge by Emily */
***Pace Difference;
PACEDIFF=&Y.UWPACE-NWPACE;
PCPACEDF=ROUND(100*PACEDIFF/&Y.UWPACE, .01);

***NWSCOREQ (Categories based on percent difference in pace);
label NWSCOREQ ="% DIFF BTW NARROW & USUAL WALKS";
IF PCPACEDF GT 15 THEN NWSCOREQ=1;
else IF 15 GE PCPACEDF GT 5 THEN NWSCOREQ=2;
else IF 5 GE PCPACEDF GE -5 THEN NWSCOREQ=3;
else IF .Z <PCPACEDF LT -5 THEN NWSCOREQ=4; /* Corrected by Ase */

```

```

        else IF DID6MNW=0 THEN NWSCOREQ=0;
        else IF DID6MNW=. THEN NWSCOREQ=.;
    end;

***** Summary Performance Scores *****;

*--- EPESEPPB;
label EPESEPPB="EPESE performance battery score 0-12";
EPESEPPB=&Y.UWSCR + CAT5CS + SBSCORE;

*---HABCPPB;
label HABCPPB="Health ABC performance score 0-4";
HABCPPB=FSBRATIO + CSRATIO + &Y.UWRATIO + NWRATIO;

*---PPB12CAT;
*THIS CODE CREATES 12 CATEGORIES OF THE CONTINUOUS HABCPPB SCORE;
label PPB12CAT="Categorical scoring of HABCPPB 0-12";
IF HABCPPB GT 3.66 THEN PPB12CAT=12;
IF 3.33 LT HABCPPB LE 3.66 THEN PPB12CAT=11;
IF 3.0 LT HABCPPB LE 3.33 THEN PPB12CAT=10;
IF 2.66 LT HABCPPB LE 3.0 THEN PPB12CAT=9;
IF 2.33 LT HABCPPB LE 2.66 THEN PPB12CAT=8;
IF 2.0 LT HABCPPB LE 2.33 THEN PPB12CAT=7;
IF 1.66 LT HABCPPB LE 2.0 THEN PPB12CAT=6;
IF 1.33 LT HABCPPB LE 1.66 THEN PPB12CAT=5;
IF 1.0 LT HABCPPB LE 1.33 THEN PPB12CAT=4;
IF 0.66 LT HABCPPB LE 1.0 THEN PPB12CAT=3;
IF 0.33 LT HABCPPB LE 0.66 THEN PPB12CAT=2;
IF 0.0 LT HABCPPB LE 0.33 THEN PPB12CAT=1;
IF HABCPPB = 0 THEN PPB12CAT=0;
%end;
format _all_;

run;
*This is all just QC code;
*****----- Formats -----
*****;
** standing balance score formats;
proc format;
    value result
        .M = "M"
        .N = "N"
        .A = "A"
        7 = "7=ref"
        9 = "9=una"
        1 = "1=<1"
        2 = "2=1-<30"
        3 = "3=30+";
    value sec_a
        .M = "M"
        .N = "N"
        .A = "A"
        1-<3 = "< 3"
        3-<10 = "3-<10"
        10-<30 = "10-<30";

```

```

value sec_b
  .M = "M"
  .N = "N"
  .A = "A"
  0 = "0"
  1-<10 = "<10"
  10-<30 = "10-<30";
value sec_e
  0 = "0"
  1-<30 = "<30"
  30 = "30";
value sec_f
  0 = "0"
  1-30 = "1-30"
  30<-60 = "31-60"
  60<-90 = "61-90";
value ratio
  0.0 = "0.0"
  .01-.33 = ".01-.33"
  .34-.67 = ".34-.67"
  .68-.99 = ".68-.99"
  1.00 = "1.00";
run;
%macro skip;
** chair stand formats;
proc format;
  value sec_c
    1-11.1 = "1-11.1"
    11.2-13.6 = "11.2-13.6"
    13.7-16.6 = "13.7-16.6"
    16.7-35 = "16.7-35"
    35<-high = ">35";
  value sec_d
    1-high = "1+";
  value pace
    0 = "0"
    0<-high = ">0";
run;

** Walking formats;
proc format;
  value walk_sec
    1.0-50.0 = "1.0-50.0";
  value walk_pace
    0 = "0"
    0<-0.46 = "<0.46"
    .47-.64 = "0.47-0.64"
    .65-.82 = "0.65-0.82"
    .83-high= "0.83+";
  value walk_pacex
    0 = "0"
    0<-high = ">0";
  value walk_steps
    1-30 = "1-30";
  value walk_ratio
    0 = "0"
    0<-0.23 = "<0.23"

```

```

    .24-.32 = "0.24-0.32"
    .33-.41 = "0.33-0.41"
    .42-high= "0.42+";
value pace_diff
  low-<0 = "<0"
    0 = "0"
  0<--<100 = ">0 - <100"
  100 = "100";
value pcpace_diff
  15<-high = "> 15"
  5<-15 = ">5 to 15"
  -5-5 = ">-5 to 5"
  low<-5 = "<-5";
value pace_ratio
  0 = "0"
  0<--<1 = ">0 to <1.0"
  1 = "1.0";
run;

** Performance formats;
proc format;
  value perf_ratio
    0 = "0"
    0<-1 = ">0";
  value HABCPPB
    0 = "0"
    0.00<-0.33 = "> 0.00 to 0.33"
    0.33<-0.66 = "> 0.33 to 0.66"
    0.66<-1.00 = "> 0.66 to 1.00"
    1.00<-1.33 = "> 1.00 to 1.33"
    1.33<-1.66 = "> 1.33 to 1.66"
    1.66<-2.00 = "> 1.66 to 2.00"
    2.00<-2.33 = "> 2.00 to 2.33"
    2.33<-2.66 = "> 2.33 to 2.66"
    2.66<-3.00 = "> 2.66 to 3.00"
    3.00<-3.33 = "> 3.00 to 3.33"
    3.33<-3.66 = "> 3.33 to 3.66"
    3.66<-high = "> 3.66 ";
run;

*****----- Frequencies and prints -----
-----*****;

proc freq data=perf_&Y;
  table sttime*stscore*sts*ststm*ts1*ts2*tstm / list missing nocum nopercent;
  table fftime*ftscores*ts1*ts2*tstm*ts2tm*stscore / list missing nocum
nopercent;
  table oltime*fttime*tr1*tr2*tr1tm*tr2tm / list missing nocum nopercent;
  format sttime fftime oltime ststm tstm ts2tm tr1tm tr2tm sec_b.;
  title2 "Standing balance - Times";
run;
proc freq data=perf_&Y;
  table fsbtime*FSBRATIO*sttime*fttime*oltime / list missing nocum nopercent;
  format sttime fftime oltime fsbtime sec_f. FSBRATIO ratio.;
  title2 "Standing balance - Summary time";
run;

```

```

/*
proc print data=perf_&Y;
  where 0 < ftime < 30 and
    (tr1=3 or 0 < tr1tm < 30 or tr2=3 or 0 < tr2tm < 30);
  id habcid;
  var source oltime ftime tr1 tr2 tr1tm tr2tm;
  title2 "One-leg stand time - Cases with tandem stand < 30 forcing one-leg
stand to be 0";
run;
*/

proc freq data=perf_&Y;
  table FTSCORE*STSCORE*sts*ststm / list missing;
  format sts result. ststm sec_a.;
  title2 "Standing Balance - Semi-tandem score";
run;
proc freq data=perf_&Y ;
  table FTSCORE*STIMElt10*ts1*tstm*ts2*ts2tm / list missing;
  table FTSCORE*STIMElt10*FT1*FT2 / missing list;
  format ts1 ts2 result. tstm ts2tm sec_b.;
  title2 "Standing Balance - Full-tandem score";
run;
proc freq data=perf_&Y;
  title2 "Standing Balance - Tandem Score";
  table SBSCORE*STSCORE*FTSCORE / list missing;
run;
proc print data=perf_&Y;
  where FTSCORE=0 and (ts1 in(2,3) or ts2 in(2,3)) and (tstm>3 or ts2tm>3);
  id Habcid;
  var source sts ststm ts1 tstm ts2 ts2tm;
  title2 "Standing Balance - Cases with Full balance set to zero because Semi
balance time < 10 sec";
run;
proc freq data=perf_&Y;
  table ABLE5CS*scs*rscs
    CAT5CS*CHR5PACE*ABLE5CS*rsec / list missing nopercnt nocum;
  format scs chair3x. rcs chair4x. CHR5PACE pace. rsec sec_c.;
  title2 "Chair Stands";
run;

*****----- Save Performance File -----
*****;
%if (&Y=Y4 or &Y=Y6 or &Y=Y11) %then %do;
proc freq data=Perf_&Y;
  where source="H";
  table DID6MUW*&hpre.4mwtm1*&hpre.4mw1*&hpre.4mwtm2 / list missing nocum
nopercnt;
  format SIXMWTM &pre.UPTM1 &pre.UPTM2 walk_sec. &Y.UWPACE walk_pace. &pre.up1
&pre.up2 walk_steps. &Y.UWRATIO walk_ratio.;
  format &hpre.4mwtm1 &hpre.4mwtm2 walk_sec.;
  title2 "6M Usual walk - Home Visits";
run;

proc freq data=Perf_&Y;
  where source="P";
  table DID6MUW*yc4mwtm1*yc4mw1*yc4mwtm2 / list missing nocum nopercnt;

```

```

format &Y.UWTM &pre.UPTM1 &pre.UPTM2 walk_sec. &Y.UWPACE walk_pace. &pre.up1
&pre.up2 walk_steps. &Y.UWRATIO walk_ratio.;
format yc4mwtm1 yc4mwtm2 walk_sec.;
title2 "6M Usual walk - Proxy";
run;
%end;
%mend skip;
%if (&Y=Y1 or &Y=Y4 or &Y=Y6 or &Y=Y10 or &Y=Y11) %then %do;
data calc.Perfm_&Y (keep = habcid
    DID6MUW SIXMWTM &Y.UWPACE &Y.UWSCR &Y.UWRATIO DID6MNW NWTIME NWPACE
    NWSCORE NWSCOREQ NWRATIO ABLE5CS CHR5PACE CAT5CS CSRATIO SBSCORE
    FSBTIME FSBRATIO EPESEPPB HABCPPB PPB12CAT);
set perf_&Y;
run;
/*proc freq data=Perf_&Y;
where source="V";
table DID&Y.UW*&pre.up1*&pre.uptm1*&pre.upr1*&pre.up2
&Y.UWTM*DID&Y.UW*&pre.uptm1*&pre.uptm2
&Y.UWSCR*&Y.UWPACE*DID&Y.UW*&Y.UWTM
&Y.UWRATIO*DID&Y.UW*&Y.UWPACE / list missing nocum nopercent;
format &Y.UWTM &pre.UPTM1 &pre.UPTM2 walk_sec. &Y.UWPACE walk_pace. &pre.up1
&pre.up2 walk_steps. &Y.UWRATIO walk_ratio.;
title2 "6M Usual walk - Clinic Visits";
run;

proc freq data=Perf_&Y;
table did6mnw*&pre.20tr1*&pre.20tr2*&pre.20tr3*&pre.20cna*DID&Y.UW
nwttime*did6mnw*&pre.20tr1*&pre.20ct1*&pre.20tr2*&pre.20ct2*&pre.20tr3*&pre.20ct3
nwscore*nwpace*did6mnw*nwttime
nwratio*nwpace
/ list missing nocum nopercent;
format nwttime &pre.20ct1 &pre.20ct2 &pre.20ct3 walk_sec. nwpace walk_pace.
nwratio pace_ratio.;
title2 "20cm narrow walk";
run;

proc freq data=Perf_&Y;
table nwscoreq*pcpacedf*DID&Y.UW*&Y.UWPACE*did6mnw*nwpace
/ list missing nocum nopercent;
format &Y.UWPACE nwpace walk_pacex. pcpacedf pcpacediff. pacediff pace_diff.;
title2 "Comparison between 6M Walk and 20cm narrow walk";
run;

proc freq data=Perf_&Y;
table EPESEPPB*&Y.UWSCR*CAT5CS*SBSCORE
HABCPPB*FSBRATIO*CSRATIO*&Y.UWRATIO*NWRATIO
PPB12CAT*HABCPPB
/ list missing nocum nopercent;
format HABCPPB HABCPPB. FSBRATIO CSRATIO &Y.UWRATIO NWRATIO perf_ratio.;
title2 "Summary Performance Scores";
run;*/
%end;

%else %do;
data calc.Perfm_&Y (keep = habcid
    ABLE5CS CHR5PACE CAT5CS CSRATIO SBSCORE

```

```
        FSBTIME FSBRATIO);
    set perf_&Y;
run;
%end;
/*proc contents data=calc.Perfm_&Y varnum;
    title2 "Perf_&Y File";
run;

proc means data=calc.Perfm_&Y n nmiss min mean max;
run;*/

%mend perfm;
%perfm;
```

Appendix X
Physical Activity Calculated Variables
Using Clinic Visit Workbook Variables

Investigator Name: Eleanor Simonsick

Step 1: assign met estimates to each activity

Met estimates derive from Ainsworth BE, Haskell WL, Leon AS, Jacobs DR, Montoye HJ, Sallis JF, Paffenbarger RS. Compendium o classification of energy costs of human physical activities. Med Sci Sports Exerc 25(1):71-80.

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
DAFSKKWK	kcal/kg/week climbing stairs	Assigns 4.0 kcal/kg/hour of stair climbing plus an additional 1.0 kcal/kg/hour carrying a load (Y4 Qaire, page 11, Q18). (Assumes 1 flight up/down takes 30 seconds.)	If DAFS7DAY=1 then DAFSKKWK = 4.0 x DAFSNUM/120 +1.0 x DAFSLOAD/120	Correction for outliers: If DAFSNUM>210 then DAFSNUM=210; if DAFSLOAD>210 then DAFSLOAD=210; If DAFS12MO≤.z and DAS7DAY≤.z then DAFSKKWK=.; if (DAFS12MO in (0,7,8)) and (DAS7DAY≤.z or DAS7DAY=8) then DAFSKKWK=0; if DAS7DAY=0 then DAFSKKWK=0; if DAFS12MO=1 and (DAS7DAY≤.z or DAS7DAY=8) then DAFSKKWK=0; if DAFSKKWK<0 then do: if DAS7DAY=1 and DAFSNUM > 0 and (DAFSLOAD≤.z or DAFSLOADK=-1) then DAFSLOAD=0; if DAS7DAY=1 and (DAFSNUM≤.z or DAFSNUMD=-1) and DAFSLOAD > 0 then DAFSNUM=DAFSLOAD
(DAEWKKWK) (DAOWKKWK)	kcal/kg/week walking for exercise or doing other walking	N/A (questions not asked separately in Year 4. Q17, page 10 combines exercise and other walking) See DATWKKWK below		

Physical Activity Calculated Variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
DATWKKWK	kcal/kg/week walking for exercise	Assigns 4.0 kcal/kg/hour walking briskly, 3.0 to walking at moderate pace, and 2.0 to strolling (Y4 Qaire, page 10, Q17 - all walking)	<p>If DAEW7DAY=1 then DATWKKWK=</p> <p>4.0 x DAEWTIME* DAEWTIM/60 if DAEWPACE=1,</p> <p>3.0 x DAEWTIME* DAEWTIM/60 if DAEWPACE=2,</p> <p>and 2.0 x DAEWTIME* DAEWTIM/60 if DAEWPACE=3</p>	<p>If DAEW12MO\leq.z and DAEW7DAY\leq.z then DATWKKWK=.;</p> <p>if (DAEW12MO in (0,7,8) and DAEW7DAY\leq.z then DATWKKWK=0;</p> <p>if DAEW7DAY=0 then DATWKKWK=0;</p> <p>if DAEW12MO=1 and DAEW7DAY\leq.z then DATWKKWK=0;</p> <p>if DATWKKWK<0 then do: if DAEWTIME > 0 and DAEWTIM > 0 and (DAEWPACE\leq.z or DAEWPACE=8) then DATWKKWK=3.0 x DAEWTIME*DAEWTIM/60 (median value at baseline; used here for consistency);</p> <p>if DAEWTIME > 0 and (DAEWTIM\leq.z or DAEWTDK=-1) then DAEWTIM=35 (median value : baseline; used here for consistency);</p> <p>if (DAEWTIME\leq.z or DAEWTMDK=-1) and DAEWTIM > 0 then DAEWTIM=4 (median value a baseline; used here for consistency)</p>
(DAHIKKWK)	kcal/kg/week high intensity exercise	N/A (although Y4 Qaire, page 12, Q19 asks about high intensity exercise, the time for only the first named activity was collected, compared to the first 4 named activities at baseline. Therefore DAHIKKWK would not have been comparable to FPHIKKWK and was not created.		
DAMCKKWK	kcal/kg/week doing heavy chores	Assigns 3.5 kcal/kg/hour doing major chores (Y4 Qaire, page 9, Q16). Note: this question combines activities listed separately at baseline. DAMCKKWK is therefore equivalent to the sum of baseline variables FPPAKKWK and FPHCKKWK.	3.5 x DAHCTIME/60	<p>If DAHC12MO\leq.z and DAHC7DAY\leq.z then DAMCKKWK=.;</p> <p>if (DAHC12MO=0 or DAHC12MO in (7,8) and (DAHC7DAY\leq.z or DAHC7DAY=0) then DAMCKKWK=0;</p> <p>if DAHC7DAY=0 then DAMCKKWK=0;</p> <p>if DAHC12MO=1 and DAHC7DAY\leq.z then DAMCKKWK=0;</p> <p>if DAMCKKWK<0 then do: if DAHC12MO=1 and DAHC7DAY=1 and (DAHCTIME\leq.z or DAHCDK=- 1) then DAHCTIME=120 (median value for non- missings at baseline; used here for consistency)</p>

Physical Activity Calculated Variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
DACWKKWK	kcal/kg/week doing child/adult care	Assigns 2.5 kcal/kg/hour doing child/adult care (Y4 Qaire, page 15, Q24)	If DAVWCURA=1 then DACWKKWK=2.5 x DAVWAHAW	if DAVWCURA=0 or DAVWCURA in (7,8) then DACWKKWK=0; if DACWKKWK<0 then do: if DAVWCURA=1 and (DAVWAHAW≤z or DAVWVK=-1) then DACWKKWK=30 (median value at baseline; used here for consistency)
DAVWKKWK	kcal/kg/week doing volunteer work	Assigns 1.5 kcal/kg/hour doing volunteer work if activity level is mainly sitting, 2.0 if sitting, some standing, 2.5 if mostly standing/walking, 3.0 if mostly walking/lifting (Y4 Qaire, page15, Q23). Adjusts for months/year worked by multiplying result by DAVWMOV/12	If DAVWCURV=1 then DAVWKKWK= 1.5 x DAVWAHVW x DAVWMOV/12 if DAVWVACT=1, 2.0 x DAVWAHVW x DAVWMOV/12 if DAVWVACT=2, 2.5 x DAVWAHVW x DAVWMOV/12 if DAVWVACT=3, 3.0 x DAVWAHVW x DAVWMOV/12 if DAVWVACT=4	if DAVWCURV=0 or DAVWCURV in (7,8) then DAVWKKWK=0; if DAVWKKWK<0 the do: if DAVWCURV=1 and DAVWAHVW > 0 and DAVWVACT≤z then DAVWACT=2; if DAVWCURV=1 and DAVWAHVW > 0 and DAVWMOV≤Z then DAVWMOV=12; if DAVWCURV=1 and DAVWAHVW≤z and DAVWMOV > 0 then DAVWAHVW=4 (median value at baseline; used here for consistency)
DAPWKKWK	kcal/kg/week doing paid work	Assigns 1.5 kcal/kg/hour doing paid work if activity level is mainly sitting, 2.0 if sitting, some standing, 2.5 if mostly standing/walking, 3.0 if mostly walking/lifting (Y4 Qaire, page 14, Q22). Adjusts for months/year worked by multiplying result by DAVWMOW/12	If DVWCURJ=1 then DAPWKKWK= 1.5 x DAVWAHWR x DAVWMOW/12 if DAVWWACT=1, 2.0 x DAVWAHWR x DAVWMOW/12 if DAVWWACT=2, 2.5 x DAVWAHWR x DAVWMOW/12 if DAVWWACT=3, 3.0 x DAVWAHWR x DAVWMOW/12 if DAVWWACT=4	If DAVWCURJ=0 or in (7,8) then DAPWKKWK=0; if DAPWKKWK<0 then do: if DAVWCURJ=1 and DAVWAHWR > 0 and DAVWWACT≤z then DAVWWACT=2; if DAVWCURJ=1 and DAVWAHWR > 0 and DAVWMOW<0 then DAVWMOW=12; if DAVWCURJ=1 and DAVWAHWR≤z and DAVWMOW > 0 then DAVWAHWR=20 (median value ata baseline; used here for consistency)

Physical Activity Calculated Variables

Step 2: create composite variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
WSKKWK	kcal/kg/week - walking + stairs	Sum of exercise walking, other walking, and stair DAimbing variables	DATWWKKWK + DAFSKKWK	Imputed version of component variables used; SAS SUM function used to sum all non-missing components
WVCKKWK	kcal/kg/week - work, vol, caregiving	Sum of paid work, volunteer work, and caregiving variables	DAPWKKWK + DAVWKKWK + DACWKKWK	Imputed version of component variables used; SAS SUM function used to sum all non-missing components
WALKTIME	minutes walking/week	Adds minutes exercise walking and other walking in past week	WALKTIME=(DAEWTIME x DAEWTIM)	If DATWKKWK=0 then WALKTIME=0;
BKTWTIME	minutes walking briskly/week	Sum of minutes brisk exercise walking plus minutes brisk other walking	If DAEWPACE=1 then BKTWTIME= DAEWTIME x DAEWTIM; If DAEWPACE>1 then BKTWTIME=0; If DAEWKKWK=0 then BKTWTIME=0;	SAS SUM function used to sum all non-missing components

Step 3: create categorical variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
WALKCAT	minutes walking/week category	Categories of time spent on any walking per week	If WALKTIME=0 then WALKCAT=0; if 0<WALKTIME<150 then WALKCAT=1; if WALKTIME ≥ 150 then WALKCAT=2	Imputed version of component variables used
BRISK90	walks briskly ≥ 90 min/week	Dichotomous variable for walking briskly ≥90 min/week vs <90 min/week	If BKTWTIME=0 then BRISK90=0; if 0<BKTWTIME<90 then BRISK90=0; if BKTWTIME ≥ 90 then BRISK90=1	Imputed version of component variables used
BRISK180	walks briskly ≥ 180 min/week	Dichotomous variable for walking briskly ≥180 min/week vs <180 min/week	If BKTWTIME=0 then BRISK180=0; if 0<BKTWTIME<180 then BRISK180=0; if BKTWTIME ≥ 180 then BRISK180=1	Imputed version of component variables used

Physical Activity Calculated Variables
Using Core Home Visit Workbook (version 1.2) Variables

Step 1: assign met estimates to each activity

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
DAFSKKWK	kcal/kg/week climbing stairs	Assigns 4.0 kcal/kg/hour of stair climbing plus an additional 1.0 kcal/kg/hour carrying a load (Revised CHV, page 9, Q16). (Assumes 1 flight up/down takes 30 seconds.)	$4.0 \times ZCFSNUM / 120 + 1.0 \times ZCFSLOAD / 120$	Correction for outliers: If $ZCFSNUM > 210$ then $ZCFSNUM = 210$; if $ZCFSLOAD > 210$ then $ZCFSLOAD = 210$; If $ZCFS12MO \leq z$ and $ZCS7DAY \leq z$ then $DAFSKKWK = .$; if $(ZCFS12MO = 0$ or in (7,8)) and $(ZCS7DAY \leq z$ or $ZCS7DAY = 8)$ then $DAFSKKWK = 0$; if $ZCS7DAY = 0$ then $DAFSKKWK = 0$; if $ZCFS12MO = 1$ and $(ZCS7DAY \leq z$ or $ZCS7DAY = 8)$ then $DAFSKKWK = 0$; if $DAFSKKWK < 0$ then do: if $ZCS7DAY = 1$ and $ZCFSNUM > 0$ and $(ZCFSLOAD \leq z$ or $ZCFSLOADK - 1)$ then $ZCFSLOAD = 0$; if $ZCS7DAY = 1$ and $(ZCFSNUM \leq z$ or $ZCFSNUMD - 1)$ and $ZCFSLOAD > 0$ then $ZCFSNUM = ZCFSLOAI$
DATWKKWK	kcal/kg/week walking for exercise and other walking	Assigns 4.0 kcal/kg/hour walking briskly, 3.0 to walking at moderate pace, and 2.0 to strolling (CHV Workbook, page 10, Q17 - all walking)	$4.0 \times ZCEWTIME \times ZCEWTIM / 60$ if $ZCEWPACE = 1$, $3.0 \times ZCEWTIME \times ZCEWTIM / 60$ if $ZCEWPACE = 2$, and $2.0 \times ZCEWTIME \times ZCEWTIM / 60$ if $ZCEWPACE = 3$	If $ZCEW12MO \leq z$ and $ZCEW7DAY \leq z$ then $DATWKKWK = .$; if $(ZCEW12MO$ in (0,7,8) and $ZCEW7DAY \leq z$ then $DATWKKWK = 0$; if $ZCEW7DAY = 0$ then $DATWKKWK = 0$; if $ZCEW12MO = 1$ and $ZCEW7DAY \leq z$ then $DATWKKWK = 0$; if $DATWKKWK < 0$ then do: if $ZCEWTIME > 0$ and $ZCEWTIM > 0$ and $(ZCEWPACE \leq z$ or $ZCEWPACE = 8)$ then $DATWKKWK = 3.0 \times ZCEWTIME \times ZCEWTIM / 60$ (median value at baseline; used here for consistency); if $ZCEWTIME > 0$ and $(ZCEWTIM \leq z$ or $ZCEWTDK = -1)$ then $ZCEWTIM = 35$ (median value at baseline; used here for consistency); if $(ZCEWTIME \leq z$ or $ZCEWTMDK = -1)$ and $ZCEWTIM > 0$ then $ZCEWTIME = 4$ (median value at baseline; used here for consistency)

Physical Activity Calculated Variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
(DAEWKKWK) (DAOWKKWK)	kcal/kg/week walking for exercise or doing other walking	N/A (questions not asked separately at home visit. Q17, page 10 combines exercise and other walking) See DATWKKWK above		
(DAHIKKWK)	kcal/kg/week high intensity exercise	N/A (question was not asked on Core Home Visit workbook)		
(DAMCKKWK)	kcal/kg/week high intensity exercise	N/A (question was not asked on Core Home Visit workbook)		
DACWKKWK	kcal/kg/week doing child/adult care	Assigns 2.5 kcal/kg/hour doing child/adult care (Revised CHV, page 11, Q18)	2.5 x ZCVWAHAW	if ZCVWCURA=0 or ZCVWCURA in (7,8) then DACWKKWK=0; if DACWKKWK<0 then do: if ZCVWCURA=1 and (ZCVWAHAW≤z or ZCVWDK=-1) then DACWKKWK=30 (median value at baseline; used here for consistency)
(DAVWKKWK)	kcal/kg/week doing volunteer work	N/A (question not asked at home visit)		
(DAPWKKWK)	kcal/kg/week doing paid work	N/A (question not asked at home visit)		

Step 2: create composite variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
WSKKWK	kcal/kg/week - walking + stairs	Sum of exercise walking, other walking, and stair climbing variables	DATWKKWK + DAFSKKWK	Imputed version of component variables used; SAS SUM function <u>not</u> used
WVCKKWK	kcal/kg/week - work, vol. caregiving	N/A (not all components asked at home visit)		
WALKTIME	minutes walking/week	Uses combination question (Q17, page 10) to calculate total time walking in past week	WALKTIME= (ZCEWTIME x ZCEWTIM)	If DAEWKKWK=0 then WALKTIME=0;
BKTWTIME	minutes walking briskly/week	Uses combination question (Q17, page 10) to calculate total time walking briskly in past week	If ZCEWPACE=1 then BKTWTIME= ZCEWTIME x ZCEWTIM; If ZCEWPACE>1 then BKTWTIME=0	If BKEWTIME is missing then BKTWTIME is missing; If DAEWKKWK=0 then BKTWTIME=0; If ZCEW12MO in (0,7,8) then BKTWTIME=0

Physical Activity Calculated Variables

Step 3: create categorical variables

Variable	General description	Detailed Description	How variable is calculated	How to handle missing or special values
WALKCAT	minutes walking/week category	Categories of time spent on any walking per week	If WALKTIME=0 then WALKCAT=0; if 0<WALKTIME<150 then WALKCAT=1; if WALKTIME ≥ 150 then WALKCAT=2	Imputed version of component variables used
BRISK90	walks briskly ≥ 90 min/week	Dichotomous variable for walking briskly ≥90 min/week vs <90 min/week	If BKTWTIME=0 then BRISK90=0; if 0<BKTWTIME<90 then BRISK90=0; if BKTWTIME ≥ 90 then BRISK90=1	Imputed version of component variables used
BRISK180	walks briskly ≥ 180 min/week	Dichotomous variable for walking briskly ≥180 min/week vs <180 min/week	If BKTWTIME=0 then BRISK180=0; if 0<BKTWTIME<180 then BRISK180=0; if BKTWTIME ≥ 180 then BRISK180=1	Imputed version of component variables used

```
OPTIONS PAGESIZE=60 LINESIZE=94 nocenter formchar='|----|+|----+=|-\<>*'
NOFMTERR;
```

```
DATA PACTSUM (KEEP=HABCID DAMCKKWK DAPWKKWK DAVWKKWK BKTWTIME BRISK180 BRISK90
DACWKKWK
```

```
          DAFSKKWK DATWKKWK WALKCAT WALKTIME WSKKWK WVCKKWK);
*changing over to use CURRENT instead of HABC2;
*   set habc4.da;
set current.y4clnvis;
  by habcid;
```

```
*PROGRAM CODE TO CREATE SUMMARY MEASURES OF PHYSICAL ACTIVITY
BASED ON CALORIC EXPENDITURE. MET ESTIMATES DERIVE FROM AINSWORTH BE,
HASKELL WL, LEON AS, JACOBS DR, MONTOYE HJ, SALLIS JF, PAFFENBARGER RS.
COMPENDIUM OF PHYSICAL ACTIVITIES: DAASSIFICATION OF ENERGY COSTS OF
HUMAN PHYSICAL ACTIVIES. MED SCI SPORTS EXERC 25(1):71-80;
```

```
*STEP 1: CALCULATION OF KCALORIES PER KILOGRAM OF BODY WEIGHT PER WEEK
EXPENDED IN EACH ACTIVITY. ASSIGNED MET VALUES AND ASSUMPTIONS FOLLOW.
FOR EASE OF COMPUTATION WHILE THE DATA IS STILL DIRTY, KCAL/KG/WK FOR
EACH ACTIVITY WILL BE CALCULATED DIRECTLY FROM MINUTES ACTIVITY PERFORMED
(OR LOADS WASHED, BAGS OF GROCERIES, FLIGHTS OF STAIRS DAIMBED). IT WILL
BE ASSUMED THAT A MISSING VALUE IN THESE FIELDS INDICATES NON-PARTICIPATION
AND A VALUE OF ZERO WILL BE ASSIGNED;
```

```
*MET ASSIGNMENTS AND OTHER DECISION RULES:
```

```
Q16 - GARDENING/YARDWORK = 3.5 (THIS HAS BEEN DOWNGRADED FROM 4.5)
Q17 - MAJOR CHORES = 3.5 (THIS HAS BEEN DOWNGRADED FROM 4.5)
SINCE THESE ARE COMBINED IN LATER ROUNDS THE SAME MET VALUE WAS ASSIGNED
Q18 - LIGHT HOUSE WORK = 2.5
Q19 - GROCERY SHOPPING = 3.5 FOR 4 BAGS PLUS 1.25 FOR THOSE CARRIED AND
1.25 FOR THOSE UNPACKED
Q20 - LAUNDRY - ASSUMING 30 MINUTES OF WORK PER LOAD THEN 1.0 PER LOAD
WASHED PLUS .5 PER LOAD PUT AWAY
Q21 - STAIRS = 4.0 WITH THE ASSUMPTION 1 FLIGHT UP/DOWN TAKES 30 SECONDS
PLUS 1.0 MET FOR EVERY FLIGHT DAIMBED WITH A LOAD
Q22 - WALKING FOR EXERCISE - 4.0 IF BRISK, 3.0 IF MODERATE, AND 2.0 IF
STROLL (DOWNGRADED TO 3.0 AND 2.0)
Q23 - WALKING FOR OTHER REASONS - 4.0 IF BRISK, 3.0 IF MODERATE, AND 2.0
IF STROLL (DOWNGRADED TO 3.0 AND 2.0)
Q24 - AEROBICS = 5.0
Q25 - WEIGHT TRAINING = 6.0 (DOWNGRADED TO 6)
Q26 - HIGH INTENSITY EXERCISE - 4.0 IF LIGHT, 6.0 IF MODERATE, 8.0 IF
VIGOROUS (DOWNGRADED FROM 7.0 AND 10.0)
Q27 - MODERATE INTENSITY EXERCISE = 3.0 FOR ALL ACTIVITIES AND INTENSITY
LEVELS (DOWNGRADED FROM 4.0)
Q28 - PAID WORK = 1.5 FOR LEVEL 1, 2.0 FOR LEVEL 2, 2.5 FOR LEVEL 3, AND
3.0 FOR LEVEL 4
Q29 - VOLUNTEER WORK = 1.5 FOR LEVEL 1, 2.0 FOR LEVEL 2, 2.5 FOR LEVEL 3, AND
3.0 FOR LEVEL 4
Q30 - CAREGIVING = 2.5;
```

```
*HERE'S THE CODE;
```

```
*MAJOR CHORES;
```

```

IF daHC12MO<=.z AND daHC7DAY<=.z THEN DAMCKKWK=.;
IF (daHC12MO=0 OR daHC12MO IN (7,8)) AND (daHC7DAY<=.z OR daHC7DAY=0) THEN
DAMCKKWK=0;
IF daHC7DAY=0 THEN DAMCKKWK=0;
IF daHC12MO=1 AND daHC7DAY<=.z THEN DAMCKKWK=0;
IF daHC7DAY=1 THEN DAMCKKWK=3.5*daHCTIME/60;
*IMPUTED MISSING CODE;
IF DAMCKKWK LT 0 THEN DO;
    IF daHC12MO=1 AND daHC7DAY=1 AND (daHCTIME<=.z OR daHCDK=-1) THEN
daHCTIME=120;
    IF daHC7DAY=1 THEN DAMCKKWK=3.5*daHCTIME/60;
END;

*CLIMBING STAIRS;
*ASSUME UP/DOWN 1 FLIGHT TAKES 30 SECONDS AND ONE ADDITIONAL MET FOR
CARRYING A LOAD;
*** Cutoff for flights of stairs - only for purpose of calculated variables ***;
*** See email from MLE dated 4/6/01 ***;
IF DAFSNUM>210 then DAFSNUM=210;
IF DAFSLOAD>210 THEN DAFSLOAD=210;

IF DAFS12MO<=.z AND DAFS7DAY<=.z THEN DAFSCKKWK=.;
IF (DAFS12MO=0 OR DAFS12MO=8 OR DAFS12MO=7) AND (dafs7day=8 OR DAFs7day<=.z)
THEN DAFSCKKWK=0;
IF DAFs7day=0 THEN DAFSCKKWK=0;
IF DAFS12MO=1 AND (dafs7day<=.z OR DAFs7day=8) THEN DAFSCKKWK=0;
IF DAFs7day=1 THEN DAFSCKKWK=(4.0*daFSNUM/120) + (1.0*daFSLOAD/120);
*IMPUTED MISSING CODE;
IF DAFSCKKWK LT 0 THEN DO;
IF DAFs7day=1 AND DAFSNUM GT 0 AND (daFSLOAD<=.z OR DAFSLOADK=-1) THEN
DAFSLOAD=0;
IF DAFs7day=1 AND (daFSNUM<=.z OR DAFSNUMD=-1) AND DAFSLOAD GT 0 THEN
DAFSNUM=daFSLOAD;
IF DAFs7day=1 THEN DAFSCKKWK=(4.0*daFSNUM/120) + (1.0*daFSLOAD/120); END;

*WALKING FOR EXERCISE;
IF DAEWPACE=1 THEN DAEWMET=4.0;
IF DAEWPACE=2 THEN DAEWMET=3.0;
IF DAEWPACE=3 THEN DAEWMET=2.0;
IF DAEW12MO<=.z AND DAEW7DAY<=.z THEN DAEWKKWK=.;
IF (DAEW12MO IN (0,7,8)) AND (DAEW7DAY<=.z OR DAEW7DAY<=.z) THEN DAEWKKWK=0;
IF DAEW7DAY=0 THEN DAEWKKWK=0;
IF DAEW12MO=1 AND DAEW7DAY<=.z THEN DAEWKKWK=0;
IF DAEW7DAY=1 THEN DAEWKKWK=DAEWMET*DAEWTIME*DAEWTIM/60;
*IMPUTED MISSING CODE;
IF DAEWKKWK LT 0 THEN DO;
IF DAEWTIME > 0 AND DAEWTIM > 0 AND (DAEWPACE<=.z OR DAEWPACE=8) THEN
DAEWMET=3.0;
IF DAEWTIME > 0 AND (DAEWTIM<=.z OR DAEWTDK=-1) THEN DAEWTIM=35;
IF (DAEWTIME<=.z OR DAEWTMDK=-1) AND DAEWTIM > 0 THEN DAEWTIME=4;
IF DAEW7DAY=1 THEN DAEWKKWK=DAEWMET*DAEWTIME*DAEWTIM/60; END;

*New variable for total walking*;
DATWKKWK=DAEWKKWK;

```

```

*WORK, VOLUNTEER, AND CAREGIVING;
IF DAVWWACT=1 THEN DAVWWMET=1.5;
IF DAVWWACT=2 THEN DAVWWMET=2.0;
IF DAVWWACT=3 THEN DAVWWMET=2.5;
IF DAVWWACT=4 THEN DAVWWMET=3.0;
IF DAVWCURJ=0 OR DAVWCURJ IN (7,8) THEN DAPWKKWK=0;
IF DAVWCURJ=1 THEN DAPWKKWK=DAVWWMET*(DAVWAHWR*DAVWMOW/12);
*IMPUTED MISSING CODE;
IF DAPWKKWK LT 0 THEN DO;
IF DAVWCURJ=1 AND DAVWAHWR > 0 AND DAVWWACT<=.z THEN DAVWWMET=2.0;
*** Change 2 ***;
IF DAVWCURJ=1 AND DAVWAHWR > 0 AND DAVWMOW<=.z THEN DAVWMOW=12;
IF DAVWCURJ=1 AND DAVWAHWR<=.z AND DAVWMOW > 0 THEN DAVWAHWR=20;
IF DAVWCURJ=1 THEN DAPWKKWK=DAVWWMET*(DAVWAHWR*DAVWMOW/12); END;

IF DAVVWACT=1 THEN DAVVWMET=1.5;
IF DAVVWACT=2 THEN DAVVWMET=2.0;
IF DAVVWACT=3 THEN DAVVWMET=2.5;
IF DAVVWACT=4 THEN DAVVWMET=3.0;
IF DAVVCURV=0 OR DAVVCURV IN (7,8) THEN DAVVKKWK=0;
IF DAVVCURV=1 THEN DAVVKKWK=DAVVWMET*(DAVVAHVW*DAVVMOV/12);
*IMPUTED MISSING CODE;
*** Change 4 ***;
IF DAVVKKWK LT 0 THEN DO;
IF DAVVCURV=1 AND DAVVAHVW > 0 AND DAVVWACT<=.z THEN DAVVWMET=2.0;
*** Change 3 ***;
IF DAVVCURV=1 AND DAVVAHVW > 0 AND DAVVMOV<=.z THEN DAVVMOV=12;
IF DAVVCURV=1 AND DAVVAHVW<=.z AND DAVVMOV > 0 THEN DAVVAHVW=4;
IF DAVVCURV=1 THEN DAVVKKWK=DAVVWMET*(DAVVAHVW*DAVVMOV/12); END;

IF DAVWCURA=0 OR DAVWCURA IN (7,8) THEN DACWKKWK=0;
IF DAVWCURA=1 THEN DACWKKWK=2.5*DAVWAHAW;
*IMPUTED MISSING CODE;
IF DACWKKWK LT 0 THEN DO;
IF DAVWCURA=1 AND (DAVWAHAW<=.z OR DAVWDK=-1) THEN DACWKKWK=30; END;

*COMPOSITE MEASURES. SINCE THE SUM FUNCTION ASSIGNS ZERO TO MISSING VALUES,
THE IMPUTED VERSION OF THE COMPONENT VARIABLES WERE USED;
*HOUSEHOLD DUTIES;
*HHKKWK=SUM(OF daPAKKWK daHCKKWK daLWKKWK daGSKKWK daLDKKWK);
*WALKING AND STAIRS;
WSKKWK=SUM(OF DAFSKKWK DATWKKWK);
*EXERCISE AND RECREATION;
*EXKKWK=SUM(OF daACKKWK daTRKKWK daHIKKWK daMIKKWK);
*WORK, VOLUNTEER, CAREGIVING;
WVCKKWK=SUM(OF DAPWKKWK DAVWKKWK DACWKKWK);
*GRAND TOTAL;
*TOTKKWK=SUM(OF HHKKWK WSKKWK EXKKWK WVCKKWK);

*CODE TO EXAMINE KCALS EXPENDED IN WALKING AND VIGOROUS EXERCISE,
SIMILAR TO PAFFENBARGER;
/*****
WSKCAL=P2WTK*WSKKWK;
EXKCAL=P2WTK*EXKKWK;
HAKCAL=SUM(OF WSKCAL EXKCAL);

```

```

*CATEGORICAL HARVARD ALUM VARIABLE;
IF HAKCAL=0 THEN HACAT=0;
IF 0 < HAKCAL < 500 THEN HACAT=1;
IF 500 LE HAKCAL < 1000 THEN HACAT=2;
IF 1000 LE HAKCAL < 1500 THEN HACAT=3;
IF 1500 LE HAKCAL < 2000 THEN HACAT=4;
IF HAKCAL GE 2000 THEN HACAT=5;
*****/

*CODE TO CALCULATE AMOUNT OF WALKING
IN MINUTES, AMOUNT OF BRISK WALKING IN MINUTES, AND ENGAGEMENT IN
INTENSE EXERCISE;

IF DATWKKWK=0 THEN WALKTIME=0;
  ELSE IF DATWKKWK>0 THEN WALKTIME=DAEWTIME*DAEWTIM;

IF WALKTIME=0 THEN WALKCAT=0;
IF 0<WALKTIME<150 THEN WALKCAT=1;
IF WALKTIME GE 150 THEN WALKCAT=2;
**Fixed code below to match original (year 1) EK 9/25/03;
IF DAEWPACE=1 THEN BKEWTIME=DAEWTIME*DAEWTIM; IF DAEWPACE>1 THEN BKEWTIME=0;
IF DATWKKWK=0 THEN BKEWTIME=0;

BKTWTIME=BKEWTIME;

IF BKTWTIME=0 THEN BRISK90=0;
IF 0<BKTWTIME<90 THEN BRISK90=0;
IF BKTWTIME GE 90 THEN BRISK90=1;

IF BKTWTIME=0 THEN BRISK180=0;
IF 0<BKTWTIME<180 THEN BRISK180=0;
IF BKTWTIME GE 180 THEN BRISK180=1;

*CODE FOR MINUTES OF HIGH INTENSITY EXERCISE. INCLUDES AEROBICS
AND WEIGHT/CIRCUIT TRAINING AND ANY HIGH INTENSITY EXERCISE;

/*****
HIGHXMIN=SUM(OF daACTIME daTRTIME daH1TIME daH2TIME daH3TIME daH4TIME);
IF daACKKWK=0 AND daTRKKWK=0 AND daH1KKWK=0 THEN HIGHXMIN=0;
IF HIGHXMIN=0 THEN HIGHX90=0;
IF 0<HIGHXMIN<90 THEN HIGHX90=0;
IF HIGHXMIN GE 90 THEN HIGHX90=1;

IF daHIKKWK=0 THEN HIGHEX=0;
IF daHIKKWK>0 THEN HIGHEX=1;
*****/

RUN;

***** CORE HOME VISIT *****;

DATA PACTSUM2 (KEEP=HABCID BKTWTIME BRISK180 BRISK90 DACWKKWK DAFSKKWK DAEWKKWK
WALKTIME WSKKWK WALKCAT);
*changing over to use CURRENT instead of HABC2;
* set habc3.zc(where=(zctype=4));
  set current.y4corehv;

```

```

    by habcid;
    if habcid=5097 then delete;
*CLIMBING STAIRS;
*ASSUME UP/DOWN 1 FLIGHT TAKES 30 SECONDS AND ONE ADDITIONAL MET FOR
CARRYING A LOAD;
*** Cutoff for flights of stairs - only for purpose of calculated variables ***;
*** See email from MLE dated 4/6/01 ***;
IF ZCFSNUM>210 then ZCFSNUM=210;
IF ZCFSLOAD>210 THEN ZCFSLOAD=210;

IF ZCFS12MO<=.z AND ZCS7DAY<=.z THEN DAFSKKWK=.;
IF (ZCFS12MO=0 OR ZCFS12MO=8 OR ZCFS12MO=7) AND (ZCS7DAY=8 OR ZCS7DAY<=.z) THEN
DAFSKKWK=0;
IF ZCS7DAY=0 THEN DAFSKKWK=0;
IF ZCFS12MO=1 AND (ZCS7DAY<=.z OR ZCS7DAY=8) THEN DAFSKKWK=0;
IF ZCS7DAY=1 THEN DAFSKKWK=(4.0*ZCFSNUM/120) + (1.0*ZCFSLOAD/120);
*IMPUTED MISSING CODE;
IF DAFSKKWK LT 0 THEN DO;
** Change 5 **;
IF ZCS7DAY=1 AND ZCFSNUM GT 0 AND (ZCFSLOAD<=.z OR ZCFSLOAD=-1) THEN ZCFSLOAD=0;
IF ZCS7DAY=1 AND (ZCFSNUM<=.z OR ZCFSNUM=-1) AND ZCFSLOAD GT 0 THEN
ZCFSNUM=ZCFSLOAD;
IF ZCS7DAY=1 THEN DAFSKKWK=(4.0*ZCFSNUM/120) + (1.0*ZCFSLOAD/120); END;

*WALKING FOR EXERCISE;
IF ZCEWPACE=1 THEN ZCEWMET=4.0;
IF ZCEWPACE=2 THEN ZCEWMET=3.0;
IF ZCEWPACE=3 THEN ZCEWMET=2.0;
IF ZCEW12MO<=.z AND ZCEW7DAY<=.z THEN DAEWKKWK=.;
IF (ZCEW12MO IN (0,7,8)) AND (ZCEW7DAY<=.z OR ZCEW7DAY=-1) THEN DAEWKKWK=0;
IF ZCEW7DAY=0 THEN DAEWKKWK=0;
IF ZCEW12MO=1 AND ZCEW7DAY<=.z THEN DAEWKKWK=0;
IF ZCEW7DAY=1 THEN DAEWKKWK=ZCEWMET*ZCEWTIME*ZCEWTIM/60;
*IMPUTED MISSING CODE;
IF DAEWKKWK LT 0 THEN DO;
IF ZCEWTIME > 0 AND ZCEWTIM > 0 AND (ZCEWPACE<=.z OR ZCEWPACE=8) THEN
ZCEWMET=3.0;
IF ZCEWTIME > 0 AND (ZCEWTIM<=.z OR ZCEWTDK=-1) THEN ZCEWTIM=35;
IF (ZCEWTIME<=.z OR ZCEWTMDK=-1) AND ZCEWTIM > 0 THEN ZCEWTIME=4;
IF ZCEW7DAY=1 THEN DAEWKKWK=ZCEWMET*ZCEWTIME*ZCEWTIM/60; END;

*CAREGIVING*;
IF ZCVWCURA=0 OR ZCVWCURA IN (7,8) THEN DACWKKWK=0;
IF ZCVWCURA=1 THEN DACWKKWK=2.5*ZCVWAHAW;
*IMPUTED MISSING CODE;
IF DACWKKWK LT 0 THEN DO;
IF ZCVWCURA=1 AND (ZCVWAHAW<=.z OR ZCVWDK=-1) THEN DACWKKWK=30; END;

*CODE TO CALCULATE AMOUNT OF WALKING
IN MINUTES, AMOUNT OF BRISK WALKING IN MINUTES, AND ENGAGEMENT IN
INTENSE EXERCISE;

IF DAEWKKWK=0 THEN WALKTIME=0;
ELSE IF DAEWKKWK>0 THEN WALKTIME=ZCEWTIME*ZCEWTIM;

IF WALKTIME=0 THEN WALKCAT=0;
IF 0<WALKTIME<150 THEN WALKCAT=1;

```

```

IF WALKTIME GE 150 THEN WALKCAT=2;

IF ZCEWPACE=1 THEN BKEWTIME=ZCEWTIME*ZCEWTIM; IF ZCEWPACE>1 THEN BKEWTIME=0;
*** One more change ***;
IF DAEWKKWK=0 THEN BKEWTIME=0;
IF ZCEW12MO=0 OR ZCEW12MO IN (7,8) THEN BKEWTIME=0;
IF ZCEW7DAY=0 THEN BKEWTIME=0;
BKTWTIME=BKEWTIME;

IF BKTWTIME=0 THEN BRISK90=0;
IF 0<BKTWTIME<90 THEN BRISK90=0;
IF BKTWTIME GE 90 THEN BRISK90=1;

IF BKTWTIME=0 THEN BRISK180=0;
IF 0<BKTWTIME<180 THEN BRISK180=0;
IF BKTWTIME GE 180 THEN BRISK180=1;

*WALKING AND STAIRS;
WSKKWK=DAFSKKWK+DAEWKKWK;

RUN;

data CALC.y4phact;
  merge /*set*/ pactsum(IN=IN1)
        pactsum2(IN=IN2 rename=(DAEWKKWK=DATWKKWK));
  by habcid;
*add this line to delete dupes caused by all ppts with core hv also being in
current.y4clnvis;
if in2 or (in1 and not in2);
***no longer needed per Emily 10/01/2002***;
  *IF IN2 THEN VISITYPE='HOME  ';
  *ELSE IF IN1 THEN VISITYPE='CLINIC';
format brisk180 brisk90 yndk. walkcat walk9x.;
LABEL
DAFSKKWK='KCAL/KG/WEEK CLIMBING STAIRS'
DATWKKWK='KCAL/KG/WEEK TOTAL WALKING'
DAPWKKWK='KCAL/KG/WEEK DOING PAID WORK'
DAVWKKWK='KCAL/KG/WEEK DOING VOLUNTEER WORK'
DACWKKWK='KCAL/KG/WEEK DOING CHILD/ADULT CARE'
DAMCKKWK='KCAL/KG/WEEK DOING MAJOR CHORES'
WSKKWK='KCAL/KG/WEEK - WALKING + STAIRS'
WVCKKWK='KCAL/KG/WEEK - WORK, VOL, CAREGIVING'
WALKTIME='MINUTES WALKING/WEEK'
WALKCAT='MINUTES WALKING/WEEK CATEGORY'
BKTWTIME='MINUTES WALKING BRISKLY/WEEK'
BRISK90='WALKS BRISKLY >= 90 MIN/WEEK'
BRISK180='WALKS BRISKLY >= 180 MIN/WEEK';
run;

```

Appendix XI
Rose Scales

Investigator Name: Catherine Carlson and Anne Newman

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
ROSEANG	Rose Angina	Summary variable from a standard questionnaire on chest pain and intermittent claudication created by Geoffrey Rose et al.	Based on questions 45-50 in the Year 4 questionnaire. See attached SAS code. definite: yes to pain or discomfort in the chest (DACSCHPN=1) and meets all of the following: (yes to gets it when walking up hill or hurrying (DACSCPUP=1) or yes to gets it when walking at an ordinary pace (DACSCPLS=1), stops or slows down or takes nitroglycerine if gets pain while walking (DACSWALK=1), pain is relieved after standing still (DACSSSTIL=1), pain is relieved in 10 minutes or less (DACSREL=2), pain occurs in the sternum (upper, middle, or lower) or left anterior chest (DACSPLOC=2 or 3). possible: yes to pain or discomfort in the chest, but does not meet all or some of the remaining criteria or has missing data for the remaining criteria. none: no to pain or discomfort in the chest (DACSCHPN=0).	= Missing if DACSCHPN is missing Don't Know, or Refused For all other component questions, missing , Don't know, and Refused are treated as No.	0= None 1= Possible 2=Definite

Rose Scales

Variable	Descriptive Title	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
ROSEIC	Rose Intermittent Claudication	Summary variable from a standard questionnaire on chest pain and intermittent claudication created by Geoffrey Rose et al.	<p>Based on question 102 from the baseline questionnaire: see attached SAS code.</p> <p>definite – yes to pain in legs while walking (DACSLGPNand meets all of the following: no to pain beginning while standing still (DACSLPSS=0), yes to getting it while walking uphill or hurrying (DACSLPUP) or yes to getting it when walking at an ordinary pace on a level surface (DACSLPLS=1), usually disappears in 10 minutes or less after standing still (DACSSTST=2), and yes to pain in calf (DACSLPCV=1).</p> <p>possible: yes to pain or discomfort in legs while walking, but either does not meet all or some of the remaining criteria or has missing data in the remaining criteria</p> <p>none: no to pain or discomfort in the legs while walking (DACSLGPN=0)</p>	<p>= Missing if DACSLGPN is missing Don't Know, or Refused</p> <p>For all other component questions, missing , Don't know, and Refused are treated as No.</p>	<p>0= None 1= Possible 2=Definite</p>

```

/*****
*Creation of Rose-Angina Variables
*Created by Cathy Carlson
*October 10, 2000
*****/
/* Updated to run on our datasets      FHarris 10/12/00 */
/*****/
* Updated for Y4 data                  7/23/02  **;
*****/

options nofmterr;
options ls=132 ps=58 formchar='|----|+|---+=|-\<>*' nocenter pageno=1 nofmterr
      mprint macrogen;

data rose(keep=habcid roseic roseang);
  set habc4.DA (keep=habcid
    DACslpup DACslpls DACslgpn DACslpss
    DACsstst DACslpcv DACscpup DACscpls
    DACschpn DACswalk DACsstil DACsrel
    DACsploc );

  if ((DACslpup eq 1) or (DACslpls eq 1)) then walkpn1=1;

  if (DACslgpn eq 0) then ROSEIC=0;
  if ((DACslgpn eq 1) and (DACslpss eq 0) and (walkpn1 eq 1)
    and (DACsstst eq 2) and (DACslpcv eq 1)) then roseic=2;
  if ((DACslgpn eq 1) and (roseic ne 2)) then roseic=1;

  if ((DACscpup eq 1) or (DACscpls eq 1)) then walkpn2=1;

  if (DACschpn eq 0) then ROSEANG=0;
  if ((DACschpn eq 1) and (walkpn2 eq 1) and (DACswalk eq 1) and (DACsstil
eq 1)
    and (DACsrel eq 1) and ((DACsploc eq 1) or (DACsploc eq 2) or
(DACsploc eq 3)))
    then roseang=2;
  if ((DACschpn eq 1) and (roseang ne 2)) then roseang=1;

  label  roseic='Rose intermittent claudication score'
        roseang='Rose angina score';

  format roseic roseang prevdzf.;
run;

data check;
  set rose;
  if roseic=.;
run;
proc print data=check;
title 'Missing values of roseic';
run;
data check;
  set rose;
  if roseang=.;
run;
proc print data=check;
title 'Missing values of roseang';

```

```
run;

data calc.rose;
  merge rose habc.ph(keep=habcid);
  by habcid;
run;
proc contents data=calc.rose;
  title 'Contents of Y4 rose angina calculated variables';
run;
proc freq data=calc.rose;
  tables roseic roseang;
run;
```

Appendix XII
Self-Reported Function Calculated Variables

Using Clinic Visit Workbook Variables

Investigator Name: Eleanor Simonsick

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASEQM	Ease walking 1/4 mile	Recode of DADWQMEZ where 3 is easiest; imputations applied for missing DADWQMEZ	If DADWQMYN in (0,7,8) then EASEQM=7-DADWQMEZ If DADWQMYN=1 then EASEQM=4-DADWQMDF	If DADWQMYN=9 and (DAMNRS>0 or DAMNRS≠22) then DADMQMYN=1; If DADWQMEZ=8 and DADWQMYN≠8 then DADWAQMEZ=2; If DADWQMDF=8 and DADWQMYN≠8 then DADWAQMDF=2; If DADWQMYN in (0,7,8) and DADWQMEZ<0 then EASEQM=4; If (DADWQMYN=0 and DADWQMEZ<0) then EASEQM=7-DADW1MEZ; if (DADWQMYN=0 and DADWQMEZ<0) and DADW1MEZ<0 and (DADW1MYN=1 or DADW1MYN=8 or DADW1MYN≤.Z) then EASEQM=4; if (DADWQMYN=0 and DADWQMEZ<0) and (DADW1MYN=0 and DADW1MEZ≤.Z) then EASEQM=5 if DADWQMYN=1 and DADWQMDF<0 then EASEQM=7-DADWQMEZ; if DADWQMYN=1 and DADWQMDF<0 and DADWQMEZ<0 then EASEQM=2; if DADWQMYN<0 and DADWQMDF>0 then EASEQM=4-DADWQMDF; if DADWQMYN<0 and DADWQMEZ>0 then EASEQM=7-DADWQMEZ; if DADWQMYN=8 and DADWQMEZ=8 then EASEQM=4; if DADWQMYN=9 and DAMNRS<0 and (DAMNRS4<0 or DAMNRS=22) then EASEQM=.	6=very easy 5=somewh at easy 4=not that easy 3=a little difficult 2=somewh at difficult 1=very difficult 0=unable to do

Self-Reported Function Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE1M	Ease walking 1 mile	Recode of DADW1MEZ where 3 is easiest; imputations applied for missing DADW1MEZ	If DADW1MYN=1 then EASE1M=0; if DADW1MEZ=3 then EASE1M=1; if DADW1MEZ=2 then EASE1M=2; if DADW1MEZ=1 then EASE1M=3	If DADW1MYN=8 and (DADW1MEZ<0 or DADW1MEZ=8) then EASE1M=1; if DADW1MYN≤Z and DADW1MEZ≤Z then EASE1M=EASEQM-4; if DADW1MYN=0 and DADW1MEZ≤Z then EASE1M=EASEQM-4; if DADWQMYN=1 then EASE1M=0; if EASEQM > 0 and (EASEQM-3 < EASE1M) and EASEQM≤3 then EASE1M=0, else If EASEQM-3<EASE1M then EASE1M=EASEQM-3 If DADW1MYN=8 and DADW1MEZ in (1,2,3) then EASE1M=4-DADW1MEZ; If DADW1MYN=0 and DADW1MEZ=8 then EASE1M=EASEQM-4;	3=very easy 2=somewhat at easy 1=not that easy 0=difficult
WKAINDEX	Walking ability index	Summary measure of self-reported walking ability.	WKAINDEX=EASEQM + EASE1M	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)
EASE1F	Ease Climbing 1 flight	Recode of DADW10EZ where 6 is easiest; imputations applied for missing DADW10EZ	EASE1F=7-DADW10EZ; If DADW10YN=1 then EASE1F=4-DADIF	If DADW10YN=9 then EASE1F=. ; If DADW10EZ=8 then DADW10EZ=2; If DADIF=8 then DADIF=2; If DADW10EZ<0 then EASE1F=7-DADW20EZ; If DADW10YN=1 and (DADIF=8 or DADIF<0) then EASE1F=2; If DADW10YN=0 and DADW10EZ<0 and DADW10EZ<0 then EASE1F=4; if EASE1F=. and DADW10YN≠9 and DADW10YN>0 and (DADW20YN=1 or DADW20YN=8 or DADW20YN≤Z) THEN EASE1F=4; if EASE1F=. and DADW20YN=0 and DADW20EZ≤Z then EASE1F=5	6=very easy 5=somewhat at easy 4=not that easy 3=a little difficult 2=somewhat at difficult 1=very difficult 0=unable to do

Self-Reported Function Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE2F	Ease Climbing 2 flights	Recode of DADW20EZ where 3 is easiest; imputations applied for missing DADW20EZ	If DADW20YN=1 then EASE2F=0; if DADW20EZ=3 then EASE2F=1; if DADW20EZ=2 then EASE2F=2; if DADW20EZ=1 then EASE2F=3	If DADW10YN=9 then EASE2F=. If (DADW20YN=8 or DADW20YN<0) and (DADW20EZ<0 or DADW20EZ=8) and EASE1F>0 then EASE2F=1; if DADW20YN≤.Z and DADW20EZ≤.Z then EASE2F=EASE1F-4; if DADW20YN=0 and (DADW20EZ≤.Z or DADW20EZ=8) and EASE1F>0 then EASE2F=EASE1F-4; if DADW10YN=1 then EASE2F=0; if .Z<EASE1F≤3 then EASE2F=0, else If EASE1F-3 < EASE2F then EASE2F=EASE1F-3	3=very easy 2=somewh at easy 1=not that easy 0=difficult
CSAINDEX	Climbing stairs ability index	Summary measure of self-reported walking ability.	CSAINDEX=EASE1F + EASE2F	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)
EASE10P	Ease lift/carry 10 lbs	Recode of DADIF10, DAEZ10LB, and DAD10AMT into a single variable, where 6 is easiest; imputations applied for missing	If DADIF10=0 or in (7,8) and DAEZ10LB ≠8 then EASE10P=7-DAEZ10LB; If DADIF10=1 then EASE10P=4-DAD10AMT	If DADIF10=0 and (DAEZ20LB=8 or DAEZ10LB<0) then EASE10P=5; otherwise If DADIF10 in (0,7,8) and DAEZ10LB=8 then EASE10P=4; If DADIF10=1 and (DAD10AMT=8 or DAD10AMT<0) then EASE10P=.; otherwise if DADIF10≤.Z and DAD10AMT>0 and DAEZ10LB<0 then EASE10P=4-DAD10AMT; if DADIF10≤.Z and DAD10AMT<0 and DAEZ10LB>0 then EASE10P=7-DAEZ10LB; if EASE10P=. and DAD20LBS ≠ 1 and DAD20LB≠8 then EASE10P=7-DAEZ20LB; if EASE10P=. and DAD20LBS=0 and DAEZ20LB=8 or ≤.Z then EASE10P=6; if EASE10P=. and DADIF10=0 and DAD20LBS=1 then EASE10P=4; if EASE10P=. and DADIF10=8 and DAD20LBS=1 then EASE10P=4; if EASE10P=. and DADIF10 in (0,8) and (DAD20LBS<0 or DAD20LBS=8) and (DAEZ20LB<0 or DAEZ20LB=8) then EASE10P=4; if EASE10P=. and DADIF10=0 and DAEZ10LB≤.Z and DAD10AMT>0 then EASE10P=4-DAD10AMT	6=very easy 5=somewh at easy 4=not that easy 3=a little difficult 2=somewh at difficult 1=very difficult 0=unable to do

Self-Reported Function Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE20P	Ease lift/carry 20 lbs	Recode of DAEZ20LB where 3 is easiest; imputations applied for missing	EASE20P=4-DAEZ20LB; if DAD20LBS=1 then EASE20P=0; if 0 le EASE10P le 3 then EASE20P=0	If DAD20LBS=8 and (DAEZ20LB=8 or DAEZ20LB<0) then EASE20p=1; If (DAD20LBS=0 or DAD20LBS<=Z) and (DAEZ20LB<=Z or DAEZ20LB=8) then EASE20P=EASE10P-4; if EASE10P-3 < EASE20P then EASE20P=EASE10P-3;	3=very easy 2=somewh at easy 1=not that easy 0=difficult
LCAINDEX	Lift/carry ability index	Summary measure of self-reported lifting/carrying ability.	LCAINDEX=EASE10P + EASE20P	Imputed version of component variables used	Ranges from 0 (unable) to 9 (very easy)
EASEUP	Ease rising from chair	Recode of DADIFSTA, DAEZSTA, and DADSTAMT into a single variable, where 6 is easiest; imputations applied for missing values	If DADIFSTA in (0,7,8) then EASEUP=7 - DAEZSTA; if DADIFSTA=1 then EASEUP=4 - DADSTAMT	If DAEZSTA=8 and DADIFSTA ne 8 then DAEZSTA=2; if DADSTAMT=8 and DADIFSTA ne 8 then DADSTAMT=2; if DADIFSTA=8 and DAEZSTA=8 then EASEUP=4; if DADIFSTA<=Z and DADSTAMT>0 and DAEZSTA<0 then EASEUP=4 - DADSTAMT; if DADIFSTA=8 and DAEZSTA<0 and DADSTAMT<0 then EASEUP=4; if DADIFSTA=0 and DAEZSTA<=Z then EASEUP=5; if DADIFSTA=0 and DAEZSTA<=Z and DADSTAMT >0 then EASEUP=4 - DADSTAMT	Ranges from 0 (least easy) to 6 (easiest)

Self-Reported Function Calculated Variables

Using Core Home Visit Workbook (version 1.2) Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASEQM	Ease walking 1/4 mile	Recode of ZCDWQMEZ where 3 is easiest; imputations applied for missing ZCDWQMEZ	If ZCDWQMYN in (0,7,8) then EASEQM=7-ZCDWQMEZ If ZCDWQMYN=1 then EASEQM=4-ZCDWQMDF	If ZCDWQMYN=9 and (ZCMNRS>0 or ZCMNRS4 in (1,2)) then ZCDMQMYN=1; If ZCDWQMEZ=8 and ZCDWQMYN≠8 then ZCDWAQMDF=2; If ZCDWQMEZ=8 and ZCDWQMYN≠8 then ZCDWAQMEZ=2; If ZCDWQMYN in (0,7,8) and ZCDWQMEZ<0 then EASEQM=4; If (ZCDWQMYN=0 and ZCDWQMEZ<0) then EASEQM=7-ZCDW1MEZ; if (ZCDWQMYN=0 and ZCDWQMEZ<0) and ZCDW1MEZ<0 and (ZCDW1MYN=1 or ZCDW1MYN=8 or ZCDW1MYN≤Z) then EASEQM=4; if (ZCDWQMYN=0 and ZCDWQMEZ<0) and (ZCDW1MYN=0 and ZCDW1MEZ≤Z) then EASEQM=5 if ZCDWQMYN=1 and ZCDWQNDF<0 then EASEQM=7-ZCDWQMEZ; if ZCDWQMYN=1 and ZCDWQNDF<0 and ZCDWQMEZ<0 then EASEQM=2; if ZCDWQMYN<0 and ZCDWQMDF>0 then EASEQM=4-ZCDWQMDF; if ZCDWQMYN<0 and ZCDWQMEZ>0 then EASEQM=7-ZCDWQMEZ; if ZCDWQMYN=8 and ZCDWQMEZ=8 then EASEQM=4; if ZCDWQMYN=9 and ZCMNRS<0 and (ZCMNRS4<0 or ZCMNRS4=8) then EASEQM=.	6=very easy 5=somewh at easy 4=not that easy 3=a little difficult 2=somewh at difficult 1=very difficult 0=unable to do

Self-Reported Function Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE1M	Ease walking 1 mile	Recode of ZCDW1MEZ where 3 is easiest; imputations applied for missing ZCDW1MEZ	If ZCDW1MYN=1 then EASE1M=0; if ZCDW1MEZ=3 then EASE1M=1; if ZCDW1MEZ=2 then EASE1M=2; if ZCDW1MEZ=1 then EASE1M=3	If ZCDW1MYN=8 and (ZCDW1MEZ<0 or ZCDW1MEZ=8) then EASE1M=1; if ZCDW1MYN≤Z and ZCDW1MEZ≤Z then EASE1M=EASEQM-4; if ZCDW1MYN=0 and ZCDW1MEZ≤Z then EASE1M=EASEQM-4; if ZCDWQMYN=1 then EASE1M=0; if EASEQM > 0 and (EASEQM-3 < EASE1M) and EASEQM≤3 then EASE1M=0, else If EASEQM-3<EASE1M then EASE1M=EASEQM-3 If ZCDW1MYN=8 and ZCDW1MEZ in (1,2,3) then EASE1M=4-ZCDW1MEZ; If ZCDW1MYN=0 and ZCDW1MEZ=8 then EASE1M=EASEQM-4;	
WKAINDEX	Walking ability index	Summary measure of self-reported walking ability.	WKAINDEX=EASEQM + EASE1M	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)
EASE1F	Ease Climbing 1 flight	Recode of ZCDW10EZ where 6 is easiest; imputations applied for missing ZCDW10EZ	EASE1F=7-ZCDW10EZ; If ZCDW10YN=1 then EASE1F=4-ZCDIF	If ZCDW10YN=9 then EASE1F=. ; If ZCDW10EZ=8 then ZCDW10EZ=2; If ZCDIF=8 then ZCDIF=2; If ZCDW10EZ<0 then EASE1F=7-ZCDW20EZ; If ZCDW10YN=1 and (ZCDIF=8 or ZCDIF<0) then EASE1F=2; If ZCDW10YN=0 and ZCDW10EZ<0 and ZCDW10EZ<0 then EASE1F=4; if EASE1F=. and ZCDW10YN≠9 and ZCDW20YN>0 and (ZCDW20YN=1 or ZCDW20YN=8 or ZCDW20YN≤Z) THEN EASE1F=4; if EASE1F=. and ZCDW20YN=0 and ZCDW20EZ≤Z then EASE1F=5	6=very easy 5=somewh at easy 4=not that easy 3=a little difficult 2=somewh at difficult 1=very difficult 0=unable to do

Self-Reported Function Calculated Variables

Variable	General Description	Detailed Description	How variable is calculated	How to handle missing or special values	Value labels
EASE2F	Ease Climbing 2 flights	Recode of ZCDW20EZ where 3 is easiest; imputations applied for missing ZCDW20EZ	If ZCDW20YN=1 then EASE2F=0; if ZCDW20EZ=3 then EASE2F=1; if ZCDW20EZ=2 then EASE2F=2; if ZCDW20EZ=1 then EASE2F=3	If ZCDW10YN=9 then EASE2F=.; If (ZCDW20YN=8 or ZCDW20YN<0) and (ZCDW20EZ<0 or ZCDW20EZ=8) and EASE1F>0 then EASE2F=1; if ZCDW20YN≤Z and ZCDW20EZ≤Z then EASE2F=EASE1F-4; if ZCDW20YN=0 and (ZCDW20EZ≤Z or ZCDW20EZ=8) and EASE1F>0 then EASE2F=EASE1F-4; if ZCDW10YN=1 then EASE2F=0; if Z<EASE1F≤3 then EASE2F=0, else If EASE1F-3 < EASE2F then EASE2F=EASE1F-3	3=very easy 2=somewh at easy 1=not that easy 0=difficult
CSAINDEX	Climbing stairs ability index	Summary measure of self-reported walking ability.	CSAINDEX=EASE1F + EASE2F	Imputed version of component variables used	This scale ranges from 0 (unable) to 9 (very easy)
EASEUP	Ease rising from chair	N/A (question not asked at home visit)			
EASE10P	Ease lift/carry 10 lbs	N/A (question not asked at home visit)			
EASE20P	Ease lift/carry 20 lbs	N/A (question not asked at home visit)			
LCAINDEX	Lift/carry ability index	N/A (question not asked at home visit)			

```

*****;
**  SRFCN_Y4.SAS **
** **
**  HABC YEAR 4 AND HOME VISIT SELF-REPORT PHYSICAL FUNCTION **
**  CALC VAR CODE **
** **
**  Creates the following variables for 6mo visit: **
**  EASEQM: ease of walking 1/4 mile **
**  EASE1M: ease of walking 1 mile **
**  EASE1F: ease of climbing 1 flight **
**  EASE2F: ease of climbing 2 flights **
**  TIREDQM: get tired walking 1/4 mile **
**  TIRED1F: get tired walking 1 mile **
**  EASE10P: ease of lifting/carrying 10 lbs **
**  EASE20P: ease of lifting/carrying 20 lbs **
**  EASEHHW: ease of doing heavy house work **
**  CSAINDEX: climbing stairs ability index **
**  WKAINDEX: walking ability index **
**  LCAINDEX: lift/carry ability index **
** **
**  Adapted from Fran Harris Y2 code: srfcn_y2.sas **
** **
**  Laura Akin 10/25/2001 **
** **
*****;
*%include '\\ilcavallo\habc\habc_sas\programs\initV8.sas';
*libname calc '\\ilcavallo\HABC\HABC_SAS\Calculated Variables\Datasets\Year
4\SAS9';
data y4; ***Y4 data***;
*substitute daf.y4clnvis so this reflects released data;
set current.y4clnvis (keep=habcid visitype DAdwqmyn DAdwqmdf DAMnrs DAdwqmez
DAdwlmyn DAdwlmez DAdw10yn DAdif
DAdw10ez DAdw20yn
DAdw20ez DAdif10 DAd10amt DAez10lb DAd20lbs
DAez20lb dadifsta daezsta dadstamt where=(visitype in
(0,10)));
*Added visitype=10 to take care of 5097,
who had a mixed visit EK 6/9/09;

* set habc4.DA(keep=habcid DAdwqmyn DAdwqmdf DAMnrs DAdwqmez
DAdwlmyn DAdwlmez DAdw10yn DAdif
DAdw10ez DAdw20yn
DAdw20ez DAdif10 DAd10amt DAez10lb DAd20lbs
DAez20lb dadifsta daezsta dadstamt);
*EASEQM: CODE FOR EASE OF WALKING 1/4 MILE
IF EASE LEVEL MISSING, EASE LEVEL WALKING 1 MILE WAS ASSIGNED. IF NO EASE LEVEL
FOR WALKING
1M WAS CODED AND DADW1MYN WAS YES, DONT KNOW, OR MISSING THEN NOT THAT EASY WAS
ASSIGNED. IF NO
EASE LEVEL FOR WALKING 1M WAS CODED AND DADW1MYN WAS NO AND DADW1MEZ WAS MISSING
THEN SOMEWHAT
EASY WAS ASSIGNED. THE 4 CASES WHO REPORTED DIFFICULTY WERE RECODED TO NOT THAT
EASY FOR QM
AND DIFFICULTY FOR 1M;
IF DADWQMYN=9 AND (DAMNRS>0 and DAMNRS ne 22) THEN DADWQMYN=1;
IF DADWQMEZ=8 and DAdwqmyn ne 8 THEN DADWQMEZ=2;
IF DADWQMDF=8 and DAdwqmyn ne 8 THEN DADWQMDF=2;

```

```

IF DADWQMYN IN (0,7,8) AND DADWQMEZ<0 THEN EASEQM=4;
  ELSE IF DADWQMYN IN (0,7,8) THEN EASEQM=7-DADWQMEZ;
IF (DADWQMYN=0 AND DADWQMEZ<0) THEN EASEQM=7-DADW1MEZ;
IF (DADWQMYN=0 AND DADWQMEZ<0) AND DADW1MEZ<0 AND (DADW1MYN=1 OR DADW1MYN=8
OR DADW1MYN<=.z) THEN EASEQM=4;
IF (DADWQMYN=0 AND DADWQMEZ<0) AND (DADW1MYN=0 AND DADW1MEZ<=.z)
THEN EASEQM=5;
IF DADWQMYN=1 THEN EASEQM=4-DADWQ MDF;

IF DADWQMYN=1 AND DADWQ MDF<0 THEN EASEQM=7-DADWQMEZ;
IF DADWQMYN=1 AND DADWQ MDF<0 AND DADWQMEZ<0 THEN EASEQM=2;
IF DADWQMYN<0 AND DADWQ MDF>0 THEN EASEQM=4-DADWQ MDF;
IF DADWQMYN<0 AND DADWQMEZ>0 THEN EASEQM=7-DADWQMEZ;

if DAdwqmyN=8 and DAdwqmez=8 then easeqm=4;
IF DADWQMYN=9 AND (DAMNRS<0 or DAMNRS=22) THEN EASEQM=.;

*EASE1M: CODE FOR EASE OF WALKING 1 MILE. AS SELF-REPORTED WALKING ABILITY IS A
KEY DEPENDENT
MEASURE IN HEALTH ABC, MISSING VALUES WERE RECODED TO NON-MISSING WHENEVER A
REASON
GUESS COULD BE MADE BASED ON OTHER AVAILABLE INFORMATION. SEE BELOW FOR CODE AND
RECODING RULES FOR MISSING VALUES;

IF DADW1MYN=1 THEN EASE1M=0;
  ELSE IF DADW1MEZ=3 THEN EASE1M=1;
  ELSE IF DADW1MEZ=2 THEN EASE1M=2;
  ELSE IF DADW1MEZ=1 THEN EASE1M=3;

*MISSING VALUE RECODES;
IF DADW1MYN=8 AND (DADW1MEZ<0 OR DADW1MEZ=8) THEN EASE1M=1;
IF DADW1MYN<=.z AND DADW1MEZ<=.z AND EASEQM=4 THEN EASE1M=0;
IF DADW1MYN<=.z AND DADW1MEZ<=.z AND EASEQM=5 THEN EASE1M=1;
IF DADW1MYN<=.z AND DADW1MEZ<=.z AND EASEQM=6 THEN EASE1M=2;
IF DADW1MYN=0 AND DADW1MEZ<=.z AND EASEQM=4 THEN EASE1M=0;
IF DADW1MYN=0 AND DADW1MEZ<=.z AND EASEQM=5 THEN EASE1M=1;
IF DADW1MYN=0 AND DADW1MEZ<=.z AND EASEQM=6 THEN EASE1M=2;
IF DADWQMYN=1 THEN EASE1M=0;

*IN SOME CASES THE AMOUNT OF EASE REPORTED FOR WALKING 1 MILE WAS GREATER THAN
THE AMOUNT OF EASE REPORTED FOR 1/4 MILE. THE CODE BELOW DOWNGRADES THE EASE
LEVEL REPORTED FOR 1M TO THAT REPORTED FOR 1/4M;
IF EASEQM GT 0 AND ((EASEQM-3) LT EASE1M) THEN DO;
  IF EASEQM LE 3 THEN EASE1M=0; ELSE
  EASE1M=EASEQM-3;
END;
if DAdwlmyn=8 and DAdwlmez in (1,2,3) then easelm=4-DAdwlmez;
if DAdwlmyn=0 and DAdwlmez=8 then easelm=easeqm-4;

*CODE TO CREATE A SUMMARY MEASURE OF SELF-REPORTED WALKING ABILITY. THIS
SCALE WILL ULTIMATELY RANGE FROM 0 TO 9, BUT AT BASELINE IT RANGES FROM
4-9, AS NO ONE HAS DIFFICULTY. THIS SCALE CAN BE CREATED IN TWO WAYS: 1) KEYING
OFF THE RESPONSE TO EASE1M AND ONLY LOOKING AT EASEQM WHEN DIFFICULTY IS
REPORTED FOR 1M OR 2) SUMMING THE RESPONSES TO THE TWO QUESTIONS. I
DID IT BOTH WAYS AND FOUND THE LATTER APPROACH SUPERIOR;

WKAINDEX=EASEQM + EASE1M;

```

```

***TIREDQM;

*IF DADWQMT2<0 THEN TIREDQM=. ;
*IF DADWQMT2=8 OR DADWQMT2=9 THEN TIREDQM=1;
*IF DADWQMT2=0 THEN TIREDQM=0;
*IF DADWQMT2=1 THEN TIREDQM=1;

*CODE FOR EASE OF CLIMBING STAIRS;

*EASE1F;
IF DADW10YN=9 THEN DO; EASE1F=. ; EASE2F=. ; END;
IF DADW10EZ=8 THEN DADW10EZ=2;
IF DADIF=8 THEN DADIF=2;
EASE1F=7-DADW10EZ;
IF DADW10EZ<0 THEN EASE1F=7-DADW20EZ;
IF DADW10YN=1 THEN EASE1F=4-DADIF;
IF DADW10YN=1 AND (DADIF=8 OR DADIF<0) THEN EASE1F=2;
IF DADW10YN=0 AND DADW10EZ<0 AND DADW20EZ<0 THEN EASE1F=4;

IF EASE1F=. AND DADW10YN NE 9 AND DADW10YN>0 AND
  (DADW20YN=1 OR DADW20YN=8 OR DADW20YN<=.z OR DADW20YN<=.z)
  THEN EASE1F=4;
IF EASE1F=. AND DADW20YN=0 AND DADW20EZ<=.z THEN EASE1F=5;

*EASE2F;
IF DADW20YN=1 THEN EASE2F=0;
  ELSE IF DADW20EZ=3 THEN EASE2F=1;
  ELSE IF DADW20EZ=2 THEN EASE2F=2;
  ELSE IF DADW20EZ=1 THEN EASE2F=3;
IF (DADW20YN=8 OR DADW20YN<0)AND (DADW20EZ<0 OR DADW20EZ=8) AND EASE1F>0 THEN
EASE2F=1;
IF DADW20YN<=.z AND DADW20EZ<=.z AND EASE1F=4 THEN EASE2F=0;
IF DADW20YN<=.z AND DADW20EZ<=.z AND EASE1F=5 THEN EASE2F=1;
IF DADW20YN<=.z AND DADW20EZ<=.z AND EASE1F=6 THEN EASE2F=2;
IF DADW20YN=0 AND (DADW20EZ<=.z OR DADW20EZ=8) AND EASE1F>0 THEN EASE2F=EASE1F-
4;
IF DADW10YN=1 THEN EASE2F=0;

IF .z<EASE1F<=3 THEN EASE2F=0; ELSE
IF (EASE1F-3) LT EASE2F THEN EASE2F=EASE1F-3;

*CODE TO CREATE A SUMMARY MEASURE OF SELF-REPORTED STAIR CLIMBING ABILITY. THIS
SCALE WILL
RANGE FROM 0 TO 9, BUT AT BASELINE IT RANGES FROM 4-9;
CSAINDEX=EASE1F + EASE2F;

*IF DADW10WX=8 OR DADW10WX=9 THEN TIRED1F=1;
*IF DADW10WX=0 THEN TIRED1F=0;
*IF DADW10WX=1 THEN TIRED1F=1;

*CODE FOR LIFTING/CARRYING;
IF DADIF10=0 AND (DAEZ10LB=8 OR DAEZ10LB<0) THEN EASE10P=5;
  ELSE IF DADIF10 IN (0,7,8) and DAEz10lb ne 8 THEN EASE10P=7-DAEZ10LB;
  else if DADIF10 in (0,7,8) and DAEz10lb=8 then ease10p=4;
IF DADIF10=1 AND (DAD10AMT=8 OR DAD10AMT<0) THEN EASE10P=. ; ELSE
IF DADIF10=1 THEN EASE10P=4 - DAD10AMT;

```

```

IF DADIF10<=.z AND DAD10AMT>0 AND DAEZ10LB<0
THEN EASE10P=4 - DAD10AMT;
IF DADIF10<=.z AND DAD10AMT<0 AND DAEZ10LB>0
THEN EASE10P=7 - DAEZ10LB;
IF EASE10P=. AND DAD20LBS NE 1 AND DAD20LBS NE 8 THEN EASE10P=7-DAEZ20LB;
IF EASE10P=. AND DAD20LBS=0 AND (DAEZ20LB LT 0 OR DAD20LBS=8) THEN EASE10P=6;
IF EASE10P=. AND DADIF10=0 AND DAD20LBS=1 THEN EASE10P=4;
IF EASE10P=. AND DADIF10=8 AND DAD20LBS=1 THEN EASE10P=4;
IF EASE10P=. AND DADIF10 IN (0,8) AND (DAD20LBS<0 OR DAD20LBS=8) AND (DAEZ20LB<0
OR DAEZ20LB=8)
    THEN EASE10P=4;
IF EASE10P=. AND DADIF10=0 AND DAEZ10LB<=.z AND DAD10AMT>0
THEN EASE10P=4 - DAD10AMT;

*IF DAEZ20LB=8 and DAd20lbs ne 8 THEN DAEZ20LB=2;
EASE20P=4-DAEZ20LB;
if DAd20lbs=8 and (DAez20lb=8 or DAez20lb<0) then ease20p=1;
if DAd20lbs=0 and DAez20lb=8 then ease20p=ease10p-4;

IF DAD20LBS=1 THEN EASE20P=0;
IF (0 LE EASE10P LE 3) THEN EASE20P=0;

IF (DAD20LBS<=.z OR DAD20LBS=0) and (DAez20lb<=.z OR DAEZ20LB=8)
AND EASE10P=4 THEN EASE20P=0;
IF (DAD20LBS=0 OR DAD20LBS<=.z) AND (DAEZ20LB<=.z OR DAEZ20LB=8)
AND EASE10P=6 THEN EASE20P=2;
IF (DAD20LBS=0 OR DAD20LBS<=.z) AND (DAEZ20LB<=.z OR DAEZ20LB=8)
AND EASE10P=5 THEN EASE20P=1;

IF (EASE10P-3) LT EASE20P THEN DO;
EASE20P=EASE10P-3; IF EASE20P<0 THEN EASE20P=0; END;

*OTHER FUNCTION MEASURES;
IF DAEZSTA=8 and DAdifsta ne 8 THEN DAEZSTA=2;
IF DADSTAMT=8 and DAdifsta ne 8 THEN DADSTAMT=2;
IF DADIFSTA IN (0,7,8) THEN EASEUP=7 - DAEZSTA;
IF DADIFSTA=1 THEN EASEUP=4 - DADSTAMT;
if DAdifsta=8 and DAezsta=8 then easeup=4;
IF DADIFSTA<=.z AND DADSTAMT>0 AND DAEZSTA<0 THEN EASEUP=4 - DADSTAMT;
IF DADIFSTA=8 AND DAEZSTA<0 AND DADSTAMT<0 THEN EASEUP=4;
IF DADIFSTA=0 AND DAEZSTA<=.z THEN EASEUP=5;
IF DADIFSTA=0 AND DAEZSTA<=.z AND DADSTAMT >0
THEN EASEUP=4 - DADSTAMT;

*CODE FOR SUMMARY INDEX;
LCAINDEX=EASE10P + EASE20P;
/*
IF DADIFHW=8 AND DADHWAMT=8 THEN EASEHHW=4;
IF DADHWAMT=8 THEN DADHWAMT=2;
IF DAEZHW=8 THEN DAEZHW=2;
IF DADIFHW IN (0,7,8) THEN EASEHHW=7 - DAEZHW;
IF DADIFHW=1 and DADHWAMT<0 THEN EASEHHW=2;
    ELSE IF DADIFHW=1 THEN EASEHHW=4 - DADHWAMT;
IF DADIFHW<=.z AND DADHWAMT GT 0 AND DAEZHW<0
THEN EASEHHW=4 - DADHWAMT;
IF DADIFHW=8 AND DADHWAMT<0 AND DAEZHW<0 THEN EASEHHW=4;
IF DADIFHW=0 AND DAEZHW<=.z THEN EASEHHW=5;

```

```

IF DADIFHW=0 AND DAEZHW<=.z AND DADHWAMT GT 0
THEN EASEHHW=4 - DADHWAMT;
*/
RUN;
proc freq data=y4;
  tables
    EASEQM*DADWQMYN*DADWQMEZ*DADWQMDF*DADW1MYN*DADW1MEZ*DAMNRS
    EASE1M*DADW1MYN*DADW1MEZ*EASEQM*DADWQMYN
    WKAINDEX*EASEQM*EASE1M
    /*TIREDQM*DADWQMT2*/
    EASE1F*DADW10YN*DADW10EZ*DADW20YN*DADW20EZ*DADIF
    EASE2F*DADW20YN*DADW20EZ*EASE1F
    CSAINDEX*EASE1F*EASE2F
    /*TIRED1F*DADW10WX*/
    EASE10P*DADIF10*DAEZ10LB*DAD10AMT*DAD20LBS*DAEZ20LB
    EASE20P*DAEZ20LB*DAD20LBS*EASE10P
    LCAINDEX*EASE10P*EASE20P
    /*EASEHHW*DADIFHW*DAEZHW*DADHWAMT*/
  / MISSING LIST;
format EASEQM DADWQMYN DADWQMEZ DADWQMDF DADW1MYN DADW1MEZ DAMNRS
    EASE1M WKAINDEX /*DADWQMT2*/
    EASE1F DADW10YN DADW10EZ DADW20YN DADW20EZ DADIF
    EASE2F CSAINDEX /*TIRED1F DADW10WX */
    EASE10P DADIF10 DAEZ10LB DAD10AMT DAD20LBS DAEZ20LB
    EASE20P LCAINDEX
    /*EASEHHW DADIFHW DAEZHW DADHWAMT*/;
  title4 'Clinic visits';
run;
data home;
*substitute current.y4corehv so this reflects released data;
set current.y4corehv (keep=habcid zcdwqmyn zcdwqmdf zcmnrs zcmnrs4 zcdwqmez
    zcdwqmt2 zcdwlmyn zcdwlmez zcdw10yn zcdif
    zcmnrs2 zcmnrs3 zcdw10ez zcdw10wx zcdw20yn
    zcdw20ez);
* set habc3.ZC(keep=habcid zctype zcdwqmyn zcdwqmdf zcmnrs zcmnrs4 zcdwqmez
    zcdwqmt2 zcdwlmyn zcdwlmez zcdw10yn zcdif
    zcmnrs2 zcmnrs3 zcdw10ez zcdw10wx zcdw20yn
    zcdw20ez);
* if zctype=4; **keep only year 4 home visits;

*EASEQM: CODE FOR EASE OF WALKING 1/4 MILE
IF EASE LEVEL MISSING, EASE LEVEL WALKING 1 MILE WAS ASSIGNED. IF NO EASE LEVEL
FOR WALKING
1M WAS CODED AND ZCDW1MYN WAS YES, DONT KNOW, OR MISSING THEN NOT THAT EASY WAS
ASSIGNED. IF NO
EASE LEVEL FOR WALKING 1M WAS CODED AND ZCDW1MYN WAS NO AND ZCDW1MEZ WAS MISSING
THEN SOMEWHAT
EASY WAS ASSIGNED. THE 4 CASES WHO REPORTED DIFFICULTY WERE RECODED TO NOT THAT
EASY FOR QM
AND DIFFICULTY FOR 1M;
IF ZCDWQMYN=9 AND (ZCMNRS>0 OR ZCMNRS4 IN(1,2)) THEN ZCDWQMYN=1;
IF ZCDWQMEZ=8 and ZCDwqmyn ne 8 THEN ZCDWQMEZ=2;
IF ZCDWQMDF=8 and ZCDwqmyn ne 8 THEN ZCDWQMDF=2;
IF ZCDWQMYN IN (0,7,8) AND ZCDWQMEZ<0 THEN EASEQM=4;
ELSE IF ZCDWQMYN IN (0,7,8) THEN EASEQM=7-ZCDWQMEZ;
IF (ZCDWQMYN=0 AND ZCDWQMEZ<0) THEN EASEQM=7-ZCDW1MEZ;
IF (ZCDWQMYN=0 AND ZCDWQMEZ<0) AND ZCDW1MEZ<0 AND (ZCDW1MYN=1 OR ZCDW1MYN=8

```

```

OR ZCDW1MYN<=.z) THEN EASEQM=4;
IF (ZCDWQMYN=0 AND ZCDWQMEZ<0) AND (ZCDW1MYN=0 AND ZCDW1MEZ<=.z)
THEN EASEQM=5;
IF ZCDWQMYN=1 THEN EASEQM=4-ZCDWQ MDF;

IF ZCDWQMYN=1 AND ZCDWQ MDF<0 THEN EASEQM=7-ZCDWQMEZ;
IF ZCDWQMYN=1 AND ZCDWQ MDF<0 AND ZCDWQMEZ<0 THEN EASEQM=2;
IF ZCDWQMYN<0 AND ZCDWQ MDF>0 THEN EASEQM=4-ZCDWQ MDF;
IF ZCDWQMYN<0 AND ZCDWQMEZ>0 THEN EASEQM=7-ZCDWQMEZ;

if ZCdwqmyn=8 and ZCdwqmez=8 then easeqm=4;
IF ZCDWQMYN=9 AND ZCMNRS<0 AND (ZCMNRS4<0 OR ZCMNRS4=8) THEN EASEQM=.;

*EASE1M: CODE FOR EASE OF WALKING 1 MILE. AS SELF-REPORTED WALKING ABILITY IS A
KEY DEPENDENT
MEASURE IN HEALTH ABC, MISSING VALUES WERE RECODED TO NON-MISSING WHENEVER A
REASON
GUESS COULD BE MADE BASED ON OTHER AVAILAZCE INFORMATION. SEE BELOW FOR CODE AND
RECODING RULES FOR MISSING VALUES;

IF ZCDW1MYN=1 THEN EASE1M=0;
ELSE IF ZCDW1MEZ=3 THEN EASE1M=1;
ELSE IF ZCDW1MEZ=2 THEN EASE1M=2;
ELSE IF ZCDW1MEZ=1 THEN EASE1M=3;

*MISSING VALUE RECODES;
IF ZCDW1MYN=8 AND (ZCDW1MEZ<0 OR ZCDW1MEZ=8) THEN EASE1M=1;
IF ZCDW1MYN<=.z AND ZCDW1MEZ<=.z AND EASEQM=4 THEN EASE1M=0;
IF ZCDW1MYN<=.z AND ZCDW1MEZ<=.z AND EASEQM=5 THEN EASE1M=1;
IF ZCDW1MYN<=.z AND ZCDW1MEZ<=.z AND EASEQM=6 THEN EASE1M=2;
IF ZCDW1MYN=0 AND ZCDW1MEZ<=.z AND EASEQM=4 THEN EASE1M=0;
IF ZCDW1MYN=0 AND ZCDW1MEZ<=.z AND EASEQM=5 THEN EASE1M=1;
IF ZCDW1MYN=0 AND ZCDW1MEZ<=.z AND EASEQM=6 THEN EASE1M=2;
IF ZCDWQMYN=1 THEN EASE1M=0;

*IN SOME CASES THE AMOUNT OF EASE REPORTED FOR WALKING 1 MILE WAS GREATER THAN
THE AMOUNT OF EASE REPORTED FOR 1/4 MILE. THE CODE BELOW DOWNGRADES THE EASE
LEVEL REPORTED FOR 1M TO THAT REPORTED FOR 1/4M;
IF EASEQM GT 0 AND ((EASEQM-3) LT EASE1M) THEN DO;
IF EASEQM LE 3 THEN EASE1M=0; ELSE
EASE1M=EASEQM-3;
END;
if ZCdwlmyn=8 and ZCdwlmez in (1,2,3) then easelm=4-ZCdwlmez;
if ZCdwlmyn=0 and ZCdwlmez=8 then easelm=easeqm-4;

*CODE TO CREATE A SUMMARY MEASURE OF SELF-REPORTED WALKING ABILITY. THIS
SCALE WILL ULTIMATELY RANGE FROM 0 TO 9, BUT AT BASELINE IT RANGES FROM
4-9, AS NO ONE HAS DIFFICULTY. THIS SCALE CAN BE CREATED IN TWO WAYS: 1) KEYING
OFF THE RESPONSE TO EASE1M AND ONLY LOOKING AT EASEQM WHEN DIFFICULTY IS
REPORTED FOR 1M OR 2) SUMMING THE RESPONSES TO THE TWO QUESTIONS. I
DID IT BOTH WAYS AND FOUND THE LATTER APPROACH SUPERIOR;

WKAINDEX=EASEQM + EASE1M;

***TIREDQM;

*IF ZCDWQMT2<0 THEN TIREDQM=.;

```

```

*IF ZCDWQMT2=8 THEN TIREDQM=1;
*IF ZCDWQMT2=0 THEN TIREDQM=0;
*IF ZCDWQMT2=1 THEN TIREDQM=1;

*CODE FOR EASE OF ZCIMBING STAIRS;

*EASE1F;
IF ZCDW10YN=9 THEN DO; EASE1F=.; EASE2F=.; END;
IF ZCDW10EZ=8 THEN ZCDW10EZ=2;
IF ZCDIF=8 THEN ZCDIF=2;
EASE1F=7-ZCDW10EZ;
IF ZCDW10EZ<0 THEN EASE1F=7-ZCDW20EZ;
IF ZCDW10YN=1 THEN EASE1F=4-ZCDIF;
IF ZCDW10YN=1 AND (ZCDIF=8 OR ZCDIF<0) THEN EASE1F=2;
IF ZCDW10YN=0 AND ZCDW10EZ<0 AND ZCDW20EZ<0 THEN EASE1F=4;

IF EASE1F=. AND ZCDW10YN NE 9 AND ZCDW10YN>0 AND
  (ZCDW20YN=1 OR ZCDW20YN=8 OR ZCDW20YN<=.z OR ZCDW20YN<=.z)
  THEN EASE1F=4;
IF EASE1F=. AND ZCDW20YN=0 AND ZCDW20EZ<=.z THEN EASE1F=5;

*EASE2F;
IF ZCDW20YN=1 THEN EASE2F=0;
  ELSE IF ZCDW20EZ=3 THEN EASE2F=1;
  ELSE IF ZCDW20EZ=2 THEN EASE2F=2;
  ELSE IF ZCDW20EZ=1 THEN EASE2F=3;
IF (ZCDW20YN=8 OR ZCDW20YN<0)AND (ZCDW20EZ<0 OR ZCDW20EZ=8) AND EASE1F>0 THEN
EASE2F=1;
IF ZCDW20YN<=.z AND ZCDW20EZ<=.z AND EASE1F=4 THEN EASE2F=0;
IF ZCDW20YN<=.z AND ZCDW20EZ<=.z AND EASE1F=5 THEN EASE2F=1;
IF ZCDW20YN<=.z AND ZCDW20EZ<=.z AND EASE1F=6 THEN EASE2F=2;
IF ZCDW20YN=0 AND (ZCDW20EZ<=.z OR ZCDW20EZ=8) AND EASE1F>0 THEN EASE2F=EASE1F-
4;
IF ZCDW10YN=1 THEN EASE2F=0;

IF .z<EASE1F<=3 THEN EASE2F=0; ELSE
IF (EASE1F-3) LT EASE2F THEN EASE2F=EASE1F-3;

*CODE TO CREATE A SUMMARY MEASURE OF SELF-REPORTED STAIR ZCIMBING ABILITY. THIS
SCALE WILL
RANGE FROM 0 TO 9, BUT AT BASELINE IT RANGES FROM 4-9;
CSAINDEX=EASE1F + EASE2F;

*IF ZCDW10WX=8 OR ZCDW10WX=9 THEN TIRED1F=1;
*IF ZCDW10WX=0 THEN TIRED1F=0;
*IF ZCDW10WX=1 THEN TIRED1F=1;

*Remove HB5097 from this dataset, as they had the clinic visit quesitonnaire;
if habcid=5097 then delete;

RUN;
proc freq data=home;
  tables
    EASEQM*ZCDWQMYN*ZCDWQMEZ*ZCDWQMDF*ZCDW1MYN*ZCDW1MEZ*ZCMNRS4
    EASE1M*ZCDW1MYN*ZCDW1MEZ*EASEQM*ZCDWQMYN
    WKAINDEX*EASEQM*EASE1M
    /*TIREDQM*ZCDWQMT2*/

```

```

EASE1F*ZCDW10YN*ZCDW10EZ*ZCDW20YN*ZCDW20EZ*ZCDIF
EASE2F*ZCDW20YN*ZCDW20EZ*EASE1F
CSAINDEX*EASE1F*EASE2F
/*TIRED1F*ZCDW10WX*/
/ MISSING LIST;
format EASEQM ZCDWQMYN ZCDWQMEZ ZCDWQMDZ ZCDW1MYN ZCDW1MEZ ZCMNRS4
EASE1M WKAINDEX ZCDWQMT2
EASE1F ZCDW10YN ZCDW10EZ ZCDW20YN ZCDW20EZ ZCDIF
EASE2F CSAINDEX /*TIRED1F*/ ZCDW10WX;
title4 'Clinic visits';
run;

data CALC.srfcn_y4(keep=habcid easeqm easelm eaself ease2f /*tiredqm tired1f*/
ease10p ease20p easeup csaindex wkaindex lcaindex);
set y4 home;
by habcid;
LABEL
EASEQM='EASE WALKING 1/4 MILE, 6=VERY EASY'
EASE1M='EASE WALKING 1 MILE, 3=VERY EASY'
WKAINDEX='WALKING ABILITY INDEX, 9=BEST'
/*TIREDQM='GETS TIRED WALKING 1/4 MILE, 1=YES'*/
EASE1F='EASE CLIMBING 1 FLIGHT, 6=VERY EASY'
EASE2F='EASE CLIMBING 2 FLIGHTS, 3=VERY EASY'
CSAINDEX='CLIMBING STAIRS ABILITY INDEX, 9=BEST'
/*TIRED1F='GETS TIRED CLIMBING 1 FLIGHT, 1=YES'*/
EASE10P='EASE LIFT/CARRY 10 LBS, 6=VERY EASY'
EASE20P='EASE LIFT/CARRY 20 LBS, 3=VERY EASY'
LCAINDEX='LIFT/CARRY ABILITY INDEX, 9=BEST'
EASEUP='EASE STANDING FROM CHAIR WITHOU USING ARMS, 6=VERY EASY';

FORMAT EASE1F EASE2F EASEQM EASE1M EASE10P EASE20P EASEUP CSAINDEX WKAINDEX
LCAINDEX SPMISS.
/*TIRED1F TIREDQM YNDK.*//;
run;
***check for duplicates***;
data dupes;
set calc.srfcn_y4;
by habcid;
if not(first.habcid and last.habcid);
run;
proc print;
title4 'Duplicates in SRFCN Y4';
run;
proc contents data=calc.srfcn_y4;
title4 'Final dataset';
run;

```

**Documentation for Reading Center Data
(Y4Read, DXRean14, BMDNotes)**

YEAR 4 READING CENTER DATA (Y4READ)	146
1. General description.....	146
2. Cross reference of dataset names with exact source	146
3. Dataset structure and contents	146
4. Condition of data	147
5. Dataset index formulation and key variable mapping.....	151
6. General strategies for manipulating and merging the data.....	151
7. Permanently lost Reading Center data	151
8. Special Missing Value Codes.....	151
Appendix I PERMANENTLY MISSING READING CENTER FILES	153
Appendix II DXA WHOLE BODY SUBREGIONS: ARTIFACTS, EXCLUSIONS AND SUBSTITUTED VALUES	154
Appendix III BMDNOTES	157

HABC Reading Center Variable List: DXA Whole Body BMD Data

HABC Reading Center Variable List: ECG Data

HABC Reading Center Variable List: Isokinetic Strength (KinCom) Data

HABC Reading Center Variable List: Peroneal Motor Nerve Conduction Data

HABC Reading Center Variable List: Vibration Perception Threshold Data

YEAR 4 READING CENTER DATA (Y4READ, DXRean14, BMDNotes) \

1. General description

The year 4 reading center file (Y4Read) contains reading center results (non-form data) of enrolled participants. Each data record contains results from the following “reading centers” (there was no reading center for isokinetic strength or neuropathy, but the data were collected outside the Health ABC data system):

Reading Center	N
DXA whole body BMD	2401
ECG	2403
Isokinetic strength (KinCom)	2096
Neuropathy (PMNC)	2295
Neuropathy (vibration perception)	2321

There are 3075 observations in the Y4Read file.

Data from Health ABC biospecimen analyses has now been moved to the BIOSPECIMENS dataset, which contains data for all years.

Please refer to the Reading Center variable lists for detailed descriptions of the variables included..

2. Cross reference of dataset names with exact source

A complete list of variable names can be found under the “Proc Contents for All Datasets” link (search under Y4READ) or by searching the Datadict file (by database). Variable names can also be found on the Reading Center Variable Lists. In some previous DXA datasets the variables DXASTID, TOTBMD, TOTBMC were ambiguous as to whether they were hip or whole body variables (in years when both were done). These variables have been dropped and new variables beginning with H or W respectively added. Thus in Year 4, when only whole body DXA was done, these variables are called WDXASTID, WTOTBMD, and WTOTBMC, respectively.

NOTE: SOME READING CENTERS USE THE SAME VARIABLE NAMES EVERY YEAR. THE ANALYST IS REMINDED TO RENAME VARIABLES BEFORE MERGING WITH OTHER YEARS AS NECESSARY.

3. Dataset structure and contents

The Y4READ file contains a single observation per participant.

Key variables:

HABCID	HABC Enrollment ID without the 2-letter prefix
SITE	HABC Clinic site: 1=Memphis; 2=Pittsburgh*
ACROSTIC	HABC Acrostic
FLAGDXA	Flag set to 1 when participant has DXA data
FLAGECG	Flag set to 1 when participant has ECG data
FLAGKC	Flag set to 1 when participant has KinCom data

Documentation date: 2010-10-01

* Must link to Health ABC participant history file (PH) to add this variable.

FLAGPMNC Flag set to 1 when participant has PMNC data
FLAGNRVB Flag set to 1 when participant has nerve vibration perception data

4. Condition of data

a. Known data errors:

There are no known problems with any reading center data, except vibration (see below) at this time.

Vibration Perception Testing

In many cases the value of D2VPTOE (which toe tested, form data) does not match VIBTOE (which toe tested, Medoc data). There were far too many such errors to try to reconcile these discrepancies. Thus, both D2VPTOE and VIBTOE are included in Y4Read.

b. Strengths and weaknesses of dataset items:

ECG

It was recently discovered that an undetected column shift had occurred in the Year 4 ECG data, resulting in misnamed variables (for example, the heart rate variable HR was nearly all 0's because it contained data intended for the variable TECH_PRB). The following variables were affected: ECGDATE, TECH_PRB, HR, QRS, AXIS, PR, and QT. Apparently, the variable QTC, which was supplied by the ECG Reading Center for Year 1 was not included in the Year 4 data. The affected columns have now all been shifted back to where they belong.

Change data, calculated by the ECG Reading Center, have been included. In some cases, this involved re-reading and sometimes re-interpretation of the Year 1 data. Year 1 (baseline) ECG data will be re-released in the near future; however, it is unlikely that any of these re-interpretations will affect the results of analyses.

Peroneal Motor Nerve Conduction There were a number of anomalies in the PMNC tracings. Although there wasn't a PMNC reading center, per se, we did have an arrangement whereby any tracings flagged by the Field Center were sent to the Eastern Virginia Medical School for review. Waveforms were flagged if any values were outside of clinically acceptable ranges (<1 mV for amplitude; <20 m/s and/or >70 m/s for velocity) or if the difference between the popliteal fossa and fibular head velocity was >10 m/s. In conjunction with the Year 11 measurement for PNMC, the Year 4 PNMC waveforms were re-read for consistency in procedures. At the time of the Year 4 exam, the take-off and peak for some waveforms was incorrectly assigned by the equipment and staff had not been instructed on how to correct these errors, as they had in Year 11. A total of 1356 waveforms were edited and of these 1344 had data that changed. Any re-read waveforms that were edited and flagged according to above criteria were reviewed by a neurologist. If the participant had a valid waveform, though did not respond to the stimulation for either the popliteal fossa or fibular head, the amplitude was a '0' value. If amplitude was '0', then the corresponding velocity is missing because velocity is a calculated value based on the stimulation ($m/s = \text{distance between proximal and distal stimulation sites} / \text{latency at the proximal stimulation site minus latency at the distal stimulation site}$). Therefore if only participants with both amplitude and velocity data are analyzed, those with a non-response to stimulation will not be included. If the neurologist determined that a participant had an invalid waveform, then the data associated with that waveform were set to missing. For popliteal fossa or fibular head velocity, 97 participants had data outside of physiologic normal ranges and were excluded for either popliteal fossa or fibular head velocity (N=47 with ≤ 1 m/s; N=50 with ≥ 80 m/s).

Additionally, 12 participants had extreme outlier values excluded for either ankle amplitude (N=3 beyond Mean \pm 4 SD), popliteal fossa or fibular head amplitude (N=1 beyond Mean \pm 4 SD), or popliteal fossa or fibular head velocity (N=8 beyond Mean \pm 4 SD). A number of files have some or all data set to missing (.U, see special missing value codes, page 7). Although PMNC=1 for these cases, there is a complete absence of usable PMNC data.

Because so many records were edited, the latency and conduction velocity variables previously released have been dropped because they no longer correspond to the corrected data and could not be calculated by hand. In addition, to make it easier to do longitudinal analysis and to flag the corrected data for anyone who already used the old data in analysis, we have changed the remaining PMNC variable names to have "Y4" prefix (see PMNC Reading Center Variable List).

DXA Whole Body Data

A minor change has been made to the correction factor for fat-free mass variables for this release (version 4.02). Rather than multiplying by 0.964, the correction has been changed to 0.946 to match the value currently being used by NHANES. DXA data for all years have been recalculated and re-released simultaneously (October 2005).

- A. The Memphis whole body array BMD variables for both the total body and subregions have been corrected due to a shift in the whole body phantom. The correction was first applied to the 12-15-00 dataset in an additive manner to only total body BMD. The correction has since been changed to multiplicative to allow application to subregions and to correspond to the correction to BMC detailed in section B below.

The correction has been applied to all Memphis whole body scans performed from 03-04-1999 through 05-30-2001, inclusive. All whole body BMD variables (total and subregion) were multiplied by 1.023.

- B. The Memphis whole body array BMC variables for both the total body and subregions have been corrected due to a shift in the whole body phantom. This corrected BMC was used to calculate the FFM minus BMC (the CTOTLEAN variable) in the corrections noted in sections E and F below.

The correction has been applied to all Memphis whole body scans performed from 03-04-1999 through 05-30-2001, inclusive. All whole body BMC variables (total and subregion) were multiplied by 1.023.

- C. The Pittsburgh whole body array BMC variables for both the total body and subregions have been corrected due to a shift in the whole body phantom. This corrected BMC was used to calculate the FFM minus BMC (the CTOTLEAN variable) in the corrections noted in sections E and F below.

The correction has been applied to all Pittsburgh whole body scans performed on 08-27-1999 through 05-30-2001, inclusive. All whole body BMC variables (total and subregion) were multiplied by 1.017.

- D. The Pittsburgh whole body array AREA variables for both the total body and subregions have been corrected due to a shift in the whole body phantom. This corrected AREA is calculated from the corrected BMC in section C and unchangedBMD.

The correction has been applied to all Pittsburgh whole body scans performed on 08-27-1999 through 05-30-2001, inclusive. All whole body AREA variables (total and subregion) were recalculated as BMC (from section C) / BMD (unchanged).

- E. It has been determined that the Pittsburgh DXA machine overestimates total mass relative to mass measured by a scale. The following correction has been applied to this dataset. It is an updated version of the correction in Memo #1542 dated 11-29-2000.

The correction was applied as follows:

1. The correction has two time intervals with multiplicative corrections to TOTFAT, TOTLEAN, and TOTMASS. In the following equations, "VAR" stands for TOTFAT, TOTLEAN, or TOTMASS:

$$\text{From 4-28-97 through 10-12-98:} \quad \text{adjusted VAR} = 0.983 * \text{VAR}$$

$$\text{From 10-13-98 through 5-30-01:} \quad \text{adjusted VAR} = 0.979 * \text{VAR}$$

2. The CTOTLEAN correction was calculated as follows:
 $\text{adjusted CTOTLEAN} = \text{adjusted TOTLEAN} - \text{WTOTBMC}$
3. The corrections in 1 and 2 were also applied to the corresponding variables in all subregions.

- F. It has been determined that for both Memphis and Pittsburgh, DXA overestimates fat-free mass compared to criterion methods. Consequently DXA also underestimates fat mass and percent fat. The following correction has been applied to this dataset. It is an updated version of the correction in Memo #1542 dated 11-29-2000.

After the above corrections in A, B, C, D, and E were applied, the following lean mass correction was applied:

1. Multiply all fat-free mass variables by 0.946:
 $\text{New fat-free mass} = 0.946 \times \text{fat-free mass}$
 $\text{New TOTLEAN} = 0.946 \times \text{TOTLEAN}$
(This is the correction that was changed in version 4.02.)
2. Recalculate all fat mass variables by subtracting the fat-free mass from the total mass of the region of interest:
 $\text{New fat mass} = \text{total mass} - \text{new fat-free mass}$
 $\text{New TOTFAT} = \text{TOTMASS} - \text{new TOTLEAN}$
3. Recalculate all lean soft tissue mass variables by subtracting the BMC from the new fat-free mass:
 $\text{New lean soft tissue mass} = \text{new fat-free mass} - \text{BMC}$
 $\text{New CTOTLEAN} = \text{New TOTLEAN} - \text{WTOTBMC}$
4. Recalculate all percent fat variables by dividing the new fat mass by the total mass:
 $\text{New percent fat} = (\text{new fat mass} / \text{total mass}) \times 100$
 $\text{New TOTPF} = (\text{new TOTFAT} / \text{TOTMASS}) \times 100$
5. The DXA scan subregions (e.g., left leg, trunk), were corrected by applying the same steps 1-4

- G. Read Appendix II: Whole Body Subregions: Artifacts, Exclusions and Substituted Values. **The whole body artifact corrections have already been applied to this dataset release by the CC.**

DXA Data for Longitudinal Analysis

Occasionally the region of interest (ROI) must be adjusted during follow up because of incident changes or inability to position the participant exactly the same for a follow-up scan as at baseline. This requires reanalysis of the preceding scans using the new ROI before meaningful change variables can be created. (For example, L1 fractures after Y1, requiring reanalysis of Y1 and subsequent scans to delete L1 from the total BMD). All change variables (see DXA Reading Center variable list) were created using the reanalyzed data, where applicable. However, the year-by-year DXA data shown in each data analysis file uses the data as originally analyzed, since this is the most correct for cross-sectional analysis. Thus, hand calculation of the change variables from Year 1-Year 4, Year 2-Year 4, or Year 3-4 using the Y4Read dataset and Y1Read, Y2Read, or Y3Read, respectively, may give slightly different values for some participants from what is supplied by the DXA Reading Center in Y4Read. Analysts are strongly encouraged to use the change variables supplied in each dataset, rather than recreating these variables.

Previous versions of the data analysis files did not give the reanalyzed values because they are not the best to use for cross-sectional analysis. However, in certain cases, particularly repeated measure analysis, it is necessary to use the individual reanalyzed values. Therefore, the DXA data are now given in two places: Y4Read, which gives the Year 4 data as analyzed in Year 4 along with the change variables utilizing any reanalysis of baseline for Year 4 ROI changes, and DXREAN14, which gives both the Y4 data (exactly as shown in Y4Read) and the reanalyzed Year 1, Year 2, and Year 3 data. The DXREAN14 dataset is in a separate downloadable file. The data from this file should only be used for these rare cases where the individual year-by-year values, rather than a calculated change variable, are used in a repeated measures model.

The DXREAN14 dataset has a single record per participant with unique variable names for each year (e.g., TOTMASS4, TOTMASS3, TOTMASS2, and TOTMASS1 for Year 4, Year 3, Year 2, and Year 1 total body mass, respectively). For most repeated measures analyses it will be necessary to reorganize the file to multiple records per participant (one record per participant-year) using an ARRAY statement. **Again, please remember that this file should only be used for repeated measures analysis. All change variables for ordinary longitudinal analyses have already been created and can be found in Y4Read.**

Isokinetic strength KinCom

Three participants had isokinetic strength testing but their electronic data were lost (see Appendix I: Permanently Missing Reading Center Files). For all three of them, the variable D3RL3 has been set to .L (see Special Missing Value Codes below) to indicate that electronic data were not available. For two of these participants (HB5229 and HB6542), there were form data backing up the electronic data. As of version 4.03, these values have been substituted for KCTMAX, KCTMEAN, KCSIDE, TRIAL, and KCSTID. All other KinCom data for these two participants have been set to .M. Note that the algorithm used to process the electronic data is a little different from the algorithm used by the KinCom software (tails of the curve are removed before calculating forces and torques), so KCTMEAN will tend to be a little higher from the form data than from the electronic data.

Nerve Vibration Threshold

For the most part, the form data for nerve vibration testing matched the electronic data exactly. For this reason, for the few participants whose electronic files were missing, the form data were substituted. Similarly, there were a handful of participants whose form data were considerably different from the electronic data. In many cases the electronic data had a date other than the date of the participant's clinic visit, and there is no record of the participant's having returned for vibration

testing. In other cases, the date was the same, but there was too much difference between the form and electronic data to ignore. After a review of the charts to determine that there were no errors in scanning, the form data were substituted for the electronic data for these participants.

5. Dataset index formulation and key variable mapping

The Y4READ file is sorted by HABCID, which is a unique identifier for each participant.

6. General strategies for manipulating and merging the data

Because the Health ABC datasets are sorted by Health ABC Enrollment ID, the HABCID variable is most useful for merging with other datasets.

7. Permanently lost Reading Center data

Every effort was made to reconcile the Reading Center databases with the clinic visit workbook data, to ensure that all measurements made on Health ABC participants that should have been sent to and analyzed by the Reading Center did reach them and became part of the Reading Center's database. As Clinic Visit Workbook data were edited, the tracking variables for Reading Center measurements were cleaned up, and multiple iterations of the reconciliation between Reading Center databases and the clinic visit data were made. Where holes in the Reading Center databases were found, the clinics were queried to try to locate the electronic or even paper version of missing measurements. Despite all of these efforts, some measurements remain irretrievable, either because they were not successfully stored by the instruments at the Field Centers, or they were not successfully transferred to the Reading Centers, or they were unrecognizable or lost at the Reading Center level. A list of participants with lost data, and the affected Reading Center database can be found in Appendix I. These participants have had the respective tracking variable values changed to .L to reflect this (see Special Missing Value Codes, below).

8. Special Missing Value Codes

SAS allows for stratification of missing values. The following missing values have been assigned:

. = 'Missing Form'
.E = 'E:Special Missing'
.L = 'L:Permanently Lost'
.M = 'M:Missing'
.T = 'T:Missing Due to Technical Problems'
.U = 'Unacceptable'

Description

. : Missing Form

Used when a value is missing because the entire form has not been entered or the participant does not exist in the database from the corresponding Reading Center.

L: Permanently Lost

Used to flag a tracking variable (e.g., D3RL3 for KinCom; D2PNDS, D2PNFHS, and D2PNPFS for PMNC) when a measurement involving a Reading Center was done, but the data either never made it to the Reading Center, or was lost at the Reading Center. A list of lost measurements, along with a brief description of what happened, can be found in Appendix I). Every effort was made to track down these data before they were declared "permanently lost" and the flag assigned.

M:Missing

Used to flag missing values when the value is required (i.e., true missing values).

T:Missing Due to Technical Problems

Used when a value is missing from the Reading Center dataset due to technical difficulties. An explanation of when this value has been assigned can be found under strengths and weaknesses of (Reading Center) dataset items for each Reading Center above.

U:Unacceptable

Used with certain Reading Center data when the data exist but cannot be used. These include:

- DXA data when the whole scan has been reviewed as unacceptable
- PMNC data when the QC process at the Eastern Virginia Medical School determined that part or all of a PMNC file could not be used for various reasons
- ECG change data when a suppression code invalidates the comparison between years of some or all aspects of a tracing (SUPPCODE=1, 2, 31, 68, 641, 711, 821, 981, 984, see ECG Reading Center Variable List)

Appendix I

PERMANENTLY MISSING READING CENTER FILES

HABC Enrollment ID#	Reading Center Database	Problem
HA2492	KinCom	Apparently done and not saved
HA1017	Neuropathy	No hardcopy; no file
HA1116	Neuropathy	No hardcopy; no file
HA2246	Neuropathy	No hardcopy; no file
HA2515	Neuropathy	No hardcopy; no file
HA2521	Neuropathy	No hardcopy; no file
HB6295	Neuropathy	No hardcopy; no file

Appendix II

DXA WHOLE BODY SUBREGIONS: ARTIFACTS, EXCLUSIONS AND SUBSTITUTED VALUES

From time to time various artifacts will appear in the subregions of the whole body, which can lead to unreliable results for the measurement of bone mineral or soft tissue. In some cases these values will need to be excluded from further analysis or they may be corrected (for example, by substituting the values from the contralateral side). This appendix summarizes some of the situations that can arise and how these exceptions are handled in preparing the DXA Reading Center.

The DXA Whole Body Reading Center dataset contains the whole body variables. See DXA Whole Body Reading Center Variable List for the variable names and labels. All variables starting with the letter “A” contain the tissue and severity codes. There can be up to two artifacts coded for each region of interest (ROI). The tissue and severity codes are provided for documentation only. **All substitutions and recalculations have already been done.** The appropriate variables contain the actual values to be used in analysis (original, substituted, or recalculated) or a missing code.

An additional file called BMDNotes is provided to give a narrative description of each artifact (see Appendix III). See the SAS Contents Procedure listing provided for this file for variable names and labels. The appropriate variables contain the narrative description, tissue code, and severity code for each artifact.

The tissue codes, severity codes, substituted and recalculated values, and missing codes are described below.

Artifacts that make the entire scan unacceptable

This is a special situation, such as motion over the entire scan, that makes it difficult to isolate the artifact to a specific subregion. In this case, there are no tissue or severity codes and all total and subregion variables are coded missing value .U.

Artifacts may affect bone, soft tissue or both

Artifacts are classified by the Reading Center according to whether they affect the bone only (code = T1), soft tissue only (code = T2), or both (code = T3). These assignments are based on empirical evidence where possible. If this is not possible, the assignment is based on theoretical grounds. The variables containing the tissue codes are assigned at the level of the whole body subregion that is affected; not at the level of the total body.

Artifacts may be of varying severity

Artifacts are also classified by the Reading Center according to their severity. The variables containing the severity codes are assigned at the level of the whole body subregion that is affected; not at the level of the total body.

Artifacts may be negligible, having no significant effect (code = S1), in which case the effect of the artifact on the measurement can be ignored.

An artifact may have a significant effect on the values measured for the subregion in which it is found (code = S2). In this case the values for the subregion have been set to missing and replaced with missing value .T.

The artifact may be more severe and have a significant effect on both the subregion and the total ROI whole body result (code = S3). In this case the values for the subregion will be set to

missing and replaced with missing value T. The values for the total region of interest will also be set to missing and replaced with missing value .T **unless** the total region of interest result can be recalculated as described in the next section below.

Artifact can be corrected by substituting values from contralateral side

Where possible, unilateral artifacts in the arm, leg, or rib (bone mineral only for rib because soft tissue artifacts in the ribs are part of the trunk ROI and there is no contralateral side to substitute) of severity grade S3 will be corrected by substituting the measurement results for the affected tissue type from the contralateral side. Substitution may be for the bone only (T1 code), soft tissue only (T2 code) or both (T3 code), except for rib as noted above. When contralateral substitution is performed, the results for the total region of interest **may** or **may not** be recalculated. The total ROIs are only recalculated if none of the component subregions are missing; e.g., there may be a leg or arm contralateral substitution, but if the pelvis is also set to .T the total will not be recalculated and will be set to the missing value .T. In practice, the total ROIs usually cannot be recalculated despite contralateral substitution, due to one or more subregions remaining missing.

If contralateral substitution results in all subregions having valid values, the recalculated total ROI values are reported in the reading center dataset. No further adjustment is needed.

Note that contralateral substitutions are only made for the purpose of recalculating the total ROI values. All subregion ROI variables retain the missing values.

An example: Suppose a scan has only one artifact: an amputation of the right leg resulting in the artifact code of T3_S3 for the right leg ROI. The reading center dataset will contain the missing value .T for the right leg ROI (this is the subregion ROI) and recalculated values for the total ROIs, in which the left leg values have been substituted for the right leg values.

Another example: Suppose a scan has only one artifact: a right hip prosthesis resulting in the artifact code of T3_S3 for the pelvis and right leg because the artifact extends into both regions. The reading center dataset will contain the missing value .T for the right leg ROI and pelvis ROI (these are the subregion ROIs) and the missing value .T for the total ROIs, because there is no contralateral substitution for the pelvis.

Actions for each specific case

Tx_S1: No action is necessary. No changes are made.

T1_S2: The bone (area, BMC, BMD) and Lean+BMC values in the affected subregion(s) are set to missing. The total results are not recalculated; the original total ROI results from the scan are retained.

T2_S2: The Fat, Lean, and Lean+BMC values in the affected subregion(s) are set to missing. The total results are not recalculated; the original total results from the scan are retained.

T3_S2: All values for the affected subregion(s) are set to missing. The total results are not recalculated; the original total results from the scan are retained.

T1_S3: The bone (area, BMC, BMD) and Lean+BMC values in the affected subregion(s) are set to missing. Values from unaffected subregions are retained for subregional analysis. To

determine if the total ROIs have been substituted and recalculated or set to missing, see section above entitled “Artifact can be corrected by substituting values from contralateral side.”

T2_S3: The Fat, Lean, Lean+BMC values in the affected subregion(s) are set to missing. Values from unaffected subregions are retained for subregional analysis. To determine if the total ROIs have been substituted and recalculated or set to missing, see section above entitled “Artifact can be corrected by substituting values from contralateral side.”

T3_S3: All values for the affected subregion(s) are set to missing. Values from unaffected subregions are retained for subregional analysis. To determine if the total ROIs have been substituted and recalculated or set to missing, see section above entitled “Artifact can be corrected by substituting values from contralateral side.”

Appendix III

BMDNOTES

General description

The BMDNotes file contains a series of 10 text variables that provide the narrative details, tissue code and severity code for each whole body artifact. This file will be of little general interest to investigators other than the DXA and Body Composition Working Group, so it was not included in Y4Read to save space. However it is available as a separate downloadable file (BMDNotes).

To link observations in BMDNotes to the corresponding observations in Y4Read, merge on the variables HABCID, WBDATE, and WBMODE.

Cross reference of dataset names with exact source

A complete list of variable names can be found under the "Proc Contents for All Datasets" link (search under BMDNotes). Variable names can also be found in the Reading Center Variable Lists.

Dataset structure and contents

The BMDNotes file contains a single observation per participant.

**DXA WHOLE BODY BMD DATA
HEALTH ABC READING CENTER VARIABLE LIST**

Variable Name	Variable Description	Variable Label	Value Label
AHEAD1	Head tissue/severity code1	Head T_S code 1	Tx_Sy, where x=1: bone only x=2: soft tissue only x =3: both y=1: negligible y=2: affects subregion only y=2: affects whole scan
AHEAD2	Head tissue/severity code2	Head T_S code 2	
ALARM1	Left arm tissue/severity code1	LARM T_S code 1	
ALARM2	Left arm tissue/severity code2	LARM T_S code 2	
ALLEG1	Left leg tissue/severity code1	LLEG T_S code 1	
ALLEG2	Left leg tissue/severity code2	LLEG T_S code 2	
ALRIBS1	Left ribs tissue/severity code1	LRIBS T_S code 1	
ALRIBS2	Left ribs tissue/severity code2	LRIBS T_S code 2	
ALSPINE1	Lumbar spine tissue/severity code1	LSPINE T_S code 1	
ALSPINE2	Lumbar spine tissue/severity code2	LSPINE T_S code 2	
APELVIS1	Pelvis tissue/severity code1	PELVIS T_S code 1	
APELVIS2	Pelvis tissue/severity code2	PELVIS T_S code 2	
ARARM1	Right arm tissue/severity code1	RARM T_S code 1	
ARARM2	Right arm tissue/severity code2	RARM T_S code 2	
ARLEG1	Right leg tissue/severity code1	RLEG T_S code 1	
ARLEG2	Right leg tissue/severity code2	RLEG T_S code 2	
ARRIBS1	Right ribs tissue/severity code1	RRIBS T_S code 1	
ARRIBS2	Right ribs tissue/severity code2	RRIBS T_S code 2	
ATSPINE1	Thoracic spine tissue/severity code1	TSPINE T_S code 1	
ATSPINE2	Thoracic spine tissue/severity code2	TSPINE T_S code 2	
CHEADLEA	Head lean	Head lean (gm)	gm
CLARMLEA	Left arm lean	Left arm lean (gm)	gm
CLLEGLEA	Left leg lean	Left leg lean (gm)	gm
CRARMLEA	Right arm lean	Right arm lean (gm)	gm
CRLEGLEA	Right leg lean	Right leg lean (gm)	gm
CTOTLEAN	Total lean	Total lean (gm)	gm
CTRNKLEA	Trunk lean	Trunk lean (gm)	gm
DCHEADx4*	Y4-x Change Head Lean (x=1, 2 or 3)	Y4-x change Head Lean (gm)	gm
DCLARMx4	Y4-x Change L Arm Lean (x=1, 2 or 3)	Y4-x change L Arm Lean (gm)	gm
DCLEANx4	Y4-x Change Total Whole Body Lean (x=1, 2 or 3)	Y4-x Change Total WBody Lean (gm)	gm
DCLLEGx4	Y4-x Change L Leg Lean (x=1, 2 or 3)	Y4-x change L Leg Lean (gm)	gm
DCRARMx4	Y4-x Change R Arm Lean (x=1, 2 or 3)	Y4-x change R Arm Lean (gm)	gm
DCRLEGx4	Y4-x Change R Leg Lean (x=1, 2 or 3)	Y4-x change R Leg Lean (gm)	gm
DCTRNX4	Y4-x Change Trunk Lean (x=1, 2 or 3)	Y4-x change Trunk Lean (gm)	gm
DFATx4	Y4-x Change Total Whole Body Fat (x=1, 2 or 3)	Y4-x Change Total WBody Fat (gm)	gm

* E.g., there are 3 variables represented by this abbreviation: DCHEAD14 (Y4-1 change Head Lean only), DCHEAD24 (Y4-2 change Head Lean only), DCHEAD34 (Y4-3 change Head Lean only)

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
DHEADAx4	Y4-x Change Head Area (x=1, 2 or 3)	Y4-x change Head Area (gm)	cm ²
DHEADCx4	Y4-x Change Head Bone Mineral Content (x=1, 2 or 3)	Y4-x change Head BMC (gm)	gm
DHEADdx4	Y4-x Change Head Bone Mineral Density (x=1, 2 or 3)	Y4-x change Head BMD (gm/cm ²)	gm/cm ²
DHEADFx4	Y4-x Change Head Fat (x=1, 2 or 3)	Y4-x change Head Fat (gm)	gm
DHEADLx4	Y4-x Change Head Fat Free Mass (x=1, 2 or 3)	Y4-x change Head FFM (gm)	gm
DLARMAx4	Y4-x Change L Arm Area (x=1, 2 or 3)	Y4-x change L Arm Area	cm ²
DLARMCx4	Y4-x Change L Arm Bone Mineral Content (x=1, 2 or 3)	Y4-x change L Arm BMC (gm)	gm
DLARMDx4	Y4-x Change L Arm Bone Mineral Density (x=1, 2 or 3)	Y4-x change L Arm BMD (gm/cm ²)	gm/cm ²
DLARMFx4	Y4-x Change L Arm Fat (x=1, 2 or 3)	Y4-x change L Arm Fat (gm)	gm
DLARMLx4	Y4-x Change L Arm Fat Free Mass (x=1, 2 or 3)	Y4-x change L Arm FFM (gm)	gm
DLEANx4	Y4-x Change Total Whole Body Fat Free Mass (x=1, 2 or 3)	Y4-x Change Total WBody FFM (gm)	gm
DLLEGAx4	Y4-x Change L Leg Area (x=1, 2 or 3)	Y4-x change L Leg Area (cm ²)	cm ²
DLLEGcx4	Y4-x Change L Leg Bone Mineral Content (x=1, 2 or 3)	Y4-x change L Leg BMC (gm)	gm
DLLEGdx4	Y4-x Change L Leg Bone Mineral Density (x=1, 2 or 3)	Y4-x change L Leg BMD (gm/cm ²)	gm/cm ²
DLLEGFx4	Y4-x Change L Leg Fat (x=1, 2 or 3)	Y4-x change L Leg Fat (gm)	gm
DLLEGLx4	Y4-x Change L Leg Fat Free Mass (x=1, 2 or 3)	Y4-x change L Leg FFM (gm)	gm
DLRIBAx4	Y4-x Change L Ribs Area (x=1, 2 or 3)	Y4-x change L Ribs Area (cm ²)	cm ²
DLRIBcx4	Y4-x Change L Ribs Bone Mineral Content (x=1, 2 or 3)	Y4-x change L Ribs BMC (gm)	gm
DLRIBdx4	Y4-x Change L Ribs Bone Mineral Density (x=1, 2 or 3)	Y4-x change L Ribs BMD (gm/cm ²)	gm/cm ²
DLSPiAx4	Y4-x Change Lumbar Spine Area (x=1, 2 or 3)	Y4-x change Lumbar Spine Area (cm ²)	cm ²
DLSPiCx4	Y4-x Change Lumbar Spine Bone Mineral Content (x=1, 2 or 3)	Y4-x change Lumbar Spine BMC (gm)	gm
DLSPiDx4	Y4-x Change Lumbar Spine Bone Mineral Density (x=1, 2 or 3)	Y4-x change Lumbar Spine BMD (gm/cm ²)	gm/cm ²
DMASSx4	Y4-x Change Total Whole Body Mass (x=1, 2 or 3)	Y4-x Change Total WBody Mass (gm)	gm
DPELVAx4	Y4-x Change Pelvis Area (x=1, 2 or 3)	Y4-x change Pelvis Area (cm ²)	cm ²
DPELVcx4	Y4-x Change Pelvis Bone Mineral Content (x=1, 2 or 3)	Y4-x change Pelvis BMC (gm)	gm
DPELVdx4	Y4-x Change Pelvis Bone Mineral Density (x=1, 2 or 3)	Y4-x change Pelvis BMD (gm/cm ²)	gm/cm ²
DPFx4	Y4-x Change Total Whole Body Percent Fat (x=1, 2 or 3)	Y4-x Change Total WBody %Fat	%

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
DRARMAx4	Y4-x Change R Arm Area (x=1, 2 or 3)	Y4-x change R Arm Area (cm2)	cm2
DRARMCx4	Y4-x Change R Arm Bone Mineral Content (x=1, 2 or 3)	Y4-x change R Arm BMC (gm)	gm
DRARMDx4	Y4-x Change R Arm Bone Mineral Density (x=1, 2 or 3)	Y4-x change R Arm BMD (gm/cm2)	gm/cm2
DRARMFx4	Y4-x Change R Arm Fat (x=1, 2 or 3)	Y4-x change R Arm Fat (gm)	gm
DRARMLx4	Y4-x Change R Arm Fat Free Mass (x=1, 2 or 3)	Y4-x change R Arm FFM (gm)	gm
DRLEGAx4	Y4-x Change R Leg Area (x=1, 2 or 3)	Y4-x change R Leg Area (cm2)	cm2
DRLEGCx4	Y4-x Change R Leg Bone Mineral Content (x=1, 2 or 3)	Y4-x change R Leg BMC (gm)	gm
DRLEGDx4	Y4-x Change R Leg Bone Mineral Density (x=1, 2 or 3)	Y4-x change R Leg BMD (gm/cm2)	gm/cm2
DRLEGFx4	Y4-x Change R Leg Fat (x=1, 2 or 3)	Y4-x change R Leg Fat (gm)	gm
DRLEGLx4	Y4-x Change R Leg Fat Free Mass (x=1, 2 or 3)	Y4-x change R Leg FFM (gm)	gm
DRRIBAx4	Y4-x Change R Ribs Area (x=1, 2 or 3)	Y4-x change R Ribs Area (cm2)	cm2
DRRIBCx4	Y4-x Change R Ribs Bone Mineral Content (x=1, 2 or 3)	Y4-x change R Ribs BMC (gm)	gm
DRRIBDx4	Y4-x Change R Ribs Bone Mineral Density (x=1, 2 or 3)	Y4-x change R Ribs BMD (gm/cm2)	gm/cm2
DTRNKCx4	Y4-x Change Trunk Bone Mineral Content (x=1, 2 or 3)	Y4-x change Trunk BMC (gm)	gm
DTRNKFx4	Y4-x Change Trunk Fat (x=1, 2 or 3)	Y4-x Change Trunk Fat (gm)	gm
DTRNKLx4	Y4-x Change Trunk Fat Free Mass (x=1, 2 or 3)	Y4-x Change Trunk FFM (gm)	gm
DTSPIAx4	Y4-x Change Thoracic Spine Area (x=1, 2 or 3)	Y4-x change Thoracic Spine Area (cm2)	cm2
DTSPICx4	Y4-x Change Thoracic Spine Bone Mineral Content (x=1, 2 or 3)	Y4-x change Thoracic Spine BMC (gm)	gm
DTSPIDx4	Y4-x Change Thoracic Spine Bone Mineral Density (x=1, 2 or 3)	Y4-x change Thoracic Spine BMD (gm/cm2)	gm/cm2
DWBAREx4	Y4-x Change Total Whole Body Area (x=1, 2 or 3)	Y4-x Change Total WBody Area	cm2
DWBBMCx4	Y4-x Change Total Whole Body Bone Mineral Content (x=1, 2 or 3)	Y4-x Change Total WBody BMC (gm)	gm
DWBBMDx4	Y4-x Change Total Whole Body Bone Mineral Density (x=1, 2 or 3)	Y4-x Change Total WBody BMD	gm/cm2
FLAGDXA	Flag set to 1 when participant has DXA whole body data	Flag set to 1 when ppt has DXA whole body data	1=Has whole body DXA data
HABCID	Health ABC Enrollment ID # without the site code	Enrollment ID without site code	
HEADAREA	Head area	Head area (cm2)	cm2
HEADBMC	Head bone mineral content	Head BMC (gm)	gm
HEADBMD	Head bone mineral density	Head BMD (gm/cm2)	gm/cm2
HEADFAT	Head fat	Head fat (gm)	gm
HEADLEAN	Head fat free mass	Head FFM (gm)	gm

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
LARMAREA	Left arm area	Left arm area (cm2)	cm2
LARMBMC	Left arm bone mineral content	Left arm BMC (gm)	gm
LARMBMD	Left arm bone mineral density	Left arm BMD (gm/cm2)	gm/cm2
LARMFAT	Left arm fat	Left arm fat (gm)	gm
LARMLEAN	Left arm fat free mass	Left arm FFM (gm)	gm
LLEGAREA	Left leg area	Left leg area (cm2)	cm2
LLEGBMC	Left leg bone mineral content	Left leg BMC (gm)	gm
LLEGBMD	Left leg bone mineral density	Left leg BMD (gm/cm2)	gm/cm2
LLEGFAT	Left leg fat	Left leg fat (gm)	gm
LLEGLEAN	Left leg fat free mass	Left leg FFM (gm)	gm
LRIBAREA	Left rib area	Left rib area (cm2)	cm2
LRIBBMC	Left rib bone mineral content	Left rib BMC (gm)	gm
LRIBBMD	Left rib bone mineral density	Left rib BMD (gm/cm2)	gm/cm2
LSPIAREA	Lumbar spine area	Lumb spine area (cm2)	cm2
LSPIBMC	Lumbar spine bone mineral content	Lumb spine BMC (gm)	gm
LSPIBMD	Lumbar spine bone mineral density	Lumb spine BMD (gm/cm2)	gm/cm2
PCHEADx4	Y4-x Percent Change Head Lean (x=1, 2 or 3)	Y4-x % change Head Lean	%
PCLARMx4	Y4-x Percent Change L Arm Lean (x=1, 2 or 3)	Y4-x % change L Arm Lean	%
PCLEANx4	Y4-x Percent Change Total Whole Body Lean (x=1, 2 or 3)	Y4-x % Change Total WBody Lean	%
PCLLEGx4	Y4-x Percent Change L Leg Lean (x=1, 2 or 3)	Y4-x % change L Leg Lean	%
PCRARMx4	Y4-x Percent Change R Arm Lean (x=1, 2 or 3)	Y4-x % change R Arm Lean	%
PCRLEGx4	Y4-x Percent Change R Leg Lean (x=1, 2 or 3)	Y4-x % change R Leg Lean	%
PCTRNX4	Y4-x Percent Change Trunk Lean (x=1, 2 or 3)	Y4-x % change Trunk Lean	%
PELVAREA	Pelvic area	Pelvic area (cm2)	cm2
PELVBMC	Pelvic bone mineral content	Pelvic BMC (gm)	gm
PELVBMD	Pelvic bone mineral density	Pelvic BMD (gm/cm2)	gm/cm2
PFATx4	Y4-x Percent Change Total Whole Body Fat (x=1, 2 or 3)	Y4-x % Change Total WBody Fat	%
PHEADAx4	Y4-x Percent Change Head Area (x=1, 2 or 3)	Y4-x % change Head Area	%
PHEADCx4	Y4-x Percent Change Head Bone Mineral Content (x=1, 2 or 3)	Y4-x % change Head BMC	%
PHEADDx4	Y4-x Percent Change Head Bone Mineral Density (x=1, 2 or 3)	Y4-x % change Head BMD	%
PHEADFx4	Y4-x Percent Change Head Fat (x=1, 2 or 3)	Y4-x % change Head Fat	%
PHEADLx4	Y4-x Percent Change Head Fat Free Mass (x=1, 2 or 3)	Y4-x % change Head FFM	%
PLARMAx4	Y4-x Percent Change L Arm Area (x=1, 2 or 3)	Y4-x % change L Arm Area	%
PLARMCx4	Y4-x Percent Change L Arm Bone Mineral Content (x=1, 2 or 3)	Y4-x % change L Arm BMC	%

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
PLARMDx4	Y4-x Percent Change L Arm Bone Mineral Density (x=1, 2 or 3)	Y4-x % change L Arm BMD	%
PLARMFx4	Y4-x Percent Change L Arm Fat (x=1, 2 or 3)	Y4-x % change L Arm Fat	%
PLARMLx4	Y4-x Percent Change L Arm Fat Free Mass (x=1, 2 or 3)	Y4-x % change L Arm FFM	%
PLEANx4	Y4-x Percent Change Total Whole Body Fat Free Mass (x=1, 2 or 3)	Y4-x % Change Total WBody FFM	%
PLLEGAx4	Y4-x Percent Change L Leg Area (x=1, 2 or 3)	Y4-x % change L Leg Area	%
PLLEGCx4	Y4-x Percent Change L Leg Bone Mineral Content (x=1, 2 or 3)	Y4-x % change L Leg BMC	%
PLLEGDx4	Y4-x Percent Change L Leg Bone Mineral Density (x=1, 2 or 3)	Y4-x % change L Leg BMD	%
PLLEGFx4	Y4-x Percent Change L Leg Fat (x=1, 2 or 3)	Y4-x % change L Leg Fat	%
PLLEGLx4	Y4-x Percent Change L Leg Fat Free Mass (x=1, 2 or 3)	Y4-x % change L Leg FFM	%
PLRIBAx4	Y4-x Percent Change L Ribs Area (x=1, 2 or 3)	Y4-x % change L Ribs Area	%
PLRIBCx4	Y4-x Percent Change L Ribs Bone Mineral Content (x=1, 2 or 3)	Y4-x % change L Ribs BMC	%
PLRIBDx4	Y4-x Percent Change L Ribs Bone Mineral Density (x=1, 2 or 3)	Y4-x % change L Ribs BMD	%
PLSPIAx4	Y4-x Percent Change Lumbar Spine Area (x=1, 2 or 3)	Y4-x % change Lumbar Spine Area	%
PLSPICx4	Y4-x Percent Change Lumbar Spine Bone Mineral Content (x=1, 2 or 3)	Y4-x % change Lumbar Spine BMC	%
PLSPIDx4	Y4-x Percent Change Lumbar Spine Bone Mineral Density (x=1, 2 or 3)	Y4-x % change Lumbar Spine BMD	%
PMASSx4	Y4-x Percent Change Total Whole Body Mass (x=1, 2 or 3)	Y4-x % Change Total WBody Mass	%
PPELVAx4	Y4-x Percent Change Pelvis Area (x=1, 2 or 3)	Y4-x % change Pelvis Area	%
PPELVCx4	Y4-x Percent Change Pelvis Bone Mineral Content (x=1, 2 or 3)	Y4-x % change Pelvis BMC	%
PPELVDx4	Y4-x Percent Change Pelvis Bone Mineral Density (x=1, 2 or 3)	Y4-x % change Pelvis BMD	%
PPFx4	Y4-x Percent Change Total Whole Body Percent Fat (x=1, 2 or 3)	Y4-x % Change Total WBody %Fat	%
PRARMAx4	Y4-x Percent Change R Arm Area (x=1, 2 or 3)	Y4-x % change R Arm Area	%
PRARMCx4	Y4-x Percent Change R Arm Bone Mineral Content (x=1, 2 or 3)	Y4-x % change R Arm BMC	%
PRARMDx4	Y4-x Percent Change R Arm Bone Mineral Density (x=1, 2 or 3)	Y4-x % change R Arm BMD	%
PRARMFx4	Y4-x Percent Change R Arm Fat (x=1, 2 or 3)	Y4-x % change R Arm Fat	%
PRARMLx4	Y4-x Percent Change R Arm Fat Free Mass (x=1, 2 or 3)	Y4-x % change R Arm FFM	%

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
PRLEGAX4	Y4-x Percent Change R Leg Area (x=1, 2 or 3)	Y4-x % change R Leg Area	%
PRLEGCx4	Y4-x Percent Change R Leg Bone Mineral Content (x=1, 2 or 3)	Y4-x % change R Leg BMC	%
PRLEGDx4	Y4-x Percent Change R Leg Bone Mineral Density (x=1, 2 or 3)	Y4-x % change R Leg BMD	%
PRLEGFx4	Y4-x Percent Change R Leg Fat (x=1, 2 or 3)	Y4-x % change R Leg Fat	%
PRLEGLx4	Y4-x Percent Change R Leg Fat Free Mass (x=1, 2 or 3)	Y4-x % change R Leg FFM	%
PRRIBAx4	Y4-x Percent Change R Ribs Area (x=1, 2 or 3)	Y4-x % change R Ribs Area	%
PRRIBCx4	Y4-x Percent Change R Ribs Bone Mineral Content (x=1, 2 or 3)	Y4-x % change R Ribs BMC	%
PRRIBDx4	Y4-x Percent Change R Ribs Bone Mineral Density (x=1, 2 or 3)	Y4-x % change R Ribs BMD	%
PTRNKCx4	Y4-x Percent Change Trunk Bone Mineral Content (x=1, 2 or 3)	Y4-x % change Trunk BMC	%
PTRNKFx4	Y4-x Percent Change Trunk Fat Free Mass (x=1, 2 or 3)	Y4-x % Change Trunk FFM	%
PTRNKLx4	Y4-x Percent Change Trunk Fat Free Mass (x=1, 2 or 3)	Y4-x % Change Trunk FFM	%
PTSPIAx4	Y4-x Percent Change Thoracic Spine Area (x=1, 2 or 3)	Y4-x % change Thoracic Spine Area	%
PTSPICx4	Y4-x Percent Change Thoracic Spine Bone Mineral Content (x=1, 2 or 3)	Y4-x % change Thoracic Spine BMC	%
PTSPIDx4	Y4-x Percent Change Thoracic Spine Bone Mineral Density (x=1, 2 or 3)	Y4-x % change Thoracic Spine BMD	%
PWBAREx4	Y4-x Percent Change Total Whole Body Area (x=1, 2 or 3)	Y4-x % Change Total WBody Area	%
PWBBMCx4	Y4-x Percent Change Total Whole Body Bone Mineral Content (x=1, 2 or 3)	Y4-x % Change Total WBody BMC	%
PWBBMDx4	Y4-x Percent Change Total Whole Body Bone Mineral Density (x=1, 2 or 3)	Y4-x % Change Total WBody BMD	%
QDR_NB	QDR serial number	QDR serial nb	%
RARMAREA	Right arm area	Right arm area (cm2)	cm ²
RARMBMC	Right arm bone mineral content	Right arm BMC (gm)	gm
RARMBMD	Right arm bone mineral density	Right arm BMD (gm/cm2)	gm/cm ²
RARMFAT	Right arm fat	Right arm fat (gm)	gm
RARMLEAN	Right arm fat free mass	Right arm FFM (gm)	gm
RLEGAREA	Right leg area	Right leg area (cm2)	cm ²
RLEGBMC	Right leg bone mineral content	Right leg BMC (gm)	gm
RLEGBMD	Right leg bone mineral density	Right leg BMD (gm/cm2)	gm/cm ²
RLEGFAT	Right leg fat (gm)	Right leg fat (gm)	gm

DXA WHOLE BODY BMD DATA

Variable Name	Variable Description	Variable Label	Value Label
RLEGLEAN	Right leg fat free mass	Right leg FFM (gm)	gm
RRIBAREA	Right rib area	Right rib area (cm2)	cm2
RRIBBMC	Right rib bone mineral content	Right rib BMC (gm)	gm
RRIBBMD	Right rib bone mineral density	Right rib BMD (gm/cm2)	gm/cm2
SITE*	Health ABC site	Field center	1 = Memphis 2= Pittsburgh
TOTFAT	Total fat (gm)	Total fat (gm)	gm
TOTLEAN	Total fat free mass	Total FFM (gm)	gm
TOTMASS	Total mass	Total mass (gm)	gm
TOTPF	Total percent fat (gm)	Total %fat (gm)	%
TRNKBMC	Trunk BMC	Trunk BMC (gm)	gm
TRNKLEAN	Trunk fat free mass	Trunk FFM (gm)	gm
TRUNKFAT	Trunk fat (gm)	Trunk fat (gm)	gm
TSPIAREA	Thoracic spine area	Thor spine area (cm2)	cm2
TSPIBMC	Thoracic spine bone mineral content	Thor spine BMC (gm)	gm
TSPIBMD	Thoracic spine bone mineral density	Thor spine BMD (gm/cm2)	gm/cm2
WBDATE	Date of exam: whole body DXA	Whole body date of exam	MMDDYY
WBDAYSx4	Y4-x # Days (x=1, 2 or 3)	Y4-x # DAYS	Days
WBFINAL	Final status: whole body	Whole body final scan status	A = acceptable N = not reviewed U = unacceptable
WBMODE	Scan mode (describes the way the scan was performed): whole body	Whole body scan mode	0 = unknown scan mode 1 = pencil beam 2 = performance mode 3 = high speed performance mode 4 = array slow 5 = array medium 6 = array fast 7 = turbo
WBSCANID	Scan ID on printout (9 character): whole body	Whole body scan ID on printout	
WBSCNPRO	Whole body scan protocol	Whole body scan protocol	
WBYRSx4	Y4-x # Years (x=1, 2 or 3)	Y4-x # YEARS	Years
WDXASTID	Staff ID #: DXA See HABC Web page for list of certified staff	Staff ID: DXA (whole body)	
WTOTAREA	Total area: whole body	Whole body total area (cm2)	cm2
WTOTBMC	Total bone mineral content: whole body	Whole body total BMC (gm)	gm
WTOTBMD	Total bone mineral density: whole body	Whole body total BMD (gm/cm2)	gm/cm2

* Must link to Health ABC participant history file (PH) to add this variable.

**HEALTH ABC READING CENTER VARIABLE LIST
RESTING ECG DATA**

Variable Name	Variable Description	Variable Label	Value Label
ABHR	Abnormal heartrate	Categories for normal HR, bradycardia (<50 bpm), tachycardia (>100 bpm)	See Y4Calc
ARRHYTH	Arrhythmia code	Rhythm irregularity	11 = frequent atrial or junctional premature beats 12 = frequent ventricular premature beats 13 = both atrial and/or junctional premature beats and ventricular premature beats 14 = wandering atrial pacemaker 15 = 12 and 14 21 = ventricular fibrillation or ventricular asystole 22 = persistent idioventricular rhythm 23 = intermittent ventricular tachycardia 24 = ventricular parasystole 31 = atrial fibrillation (persistent) 32 = atrial flutter (persistent) 33 = intermittent atrial fibrillation 34 = intermittent atrial flutter 41 = supraventricular rhythm persistent. QRS duration < 0.12 sec; and absent P-waves or presence of abnormal P-waves (inverted or flat in aVF); and regular rhythm 42 = Supraventricular tachycardia intermittent. 3 consecutive atrial or junctional premature beats occurring at a rate ≥ 100. 51 = sinoatrial arrest 52 = sinoatrial block 61 = A-V dissociation w/ ventricular pacemaker (w/out capture) 62 = A-V dissociation w/ ventricular pacemaker (w/ capture) 63 = A-V dissociation w/atrial pacemaker (w/out capture) 64 = A-V dissociation w/atrial pacemaker (w/ capture) 7 = sinus tachycardia 8 = sinus bradycardia 9 = other arrhythmias. Heart rate may be recorded as a continuous variable.
ARYTHTYP	Type of arrhythmia code	Combines ARRHYTH=31 and 32; 33 and 34; keeps 8 and 12 separate, and combines all other ARRHYTH codes	See Y4Calc

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
AV_DEF	Atrioventricular conduction defect	Atrioventricular conduction defect code	1 = complete (third degree) A-V block 21 = Mobitz type II 22 = partial (second degree) A-V block 23 = Wenckebach's phenomenon 3 = long P-R (P-Q) interval ≥ 0.22 sec 41 = Wolff-Parkinson-White pattern (WPW) 42 = WPW pattern, intermittent 5 = short P-R interval, P-R interval < 0.12 sec 6 = intermittent aberrant atrioventricular conduction 8=pacemaker
AVDEFTYP	Type of atrioventricular conduction defect	Combines AV_DEF=41 and 42; 22, 23, and 6. Keeps 1, 21, 3, and 5 separate.	See Y4Calc
AXIS	QRS axis	QRS axis	+ or – degrees
AXISAB	Axis abnormality	Categories for axis < -45 deg, $-45 \leq \text{axis} \leq 120$, $\text{axis} > 120$ degrees	See Y4Calc
C_ANT	Is the anterior lead group codeable?	Is the anterior lead group codeable?	0= No 1= Yes
C_INF	Is the inferior lead group codeable?	Is the inferior lead group codeable?	0= No 1= Yes
C_LAT	Is the lateral lead group codeable?	Is the lateral lead group codeable?	0= No 1= Yes
DCHG_ANT	Year 1-4 change in STD_ANT	ST Depression change anterior	0 = No significant change 1 = 1 grade progressions (worsening) 2 = 2 grade progressions (worsening) -1 = 1 grade regression (improvement) -2 = 2 grade regression (improvement)
DCHG_INF	Year 1-4 change in STD_INF	ST Depression change inferior	
DCHG_LAT	Year 1-4 change in STD_LAT	ST Depression change lateral	
ECG_DATE	ECG date and time	ECG date and time	MM/DD/YY HH:MM
ECG1DATE	Date/time of 1st (baseline) ECG	ECG1 date	MM/DD/YY HH:MM
ECG2DATE	Date/time of 2nd (Year 4 ECG)	ECG2 date	MM/DD/YY HH:MM
ECHG_ANT	Year 1-4 change in STL_ANL	ST Elevation change anterior	0 = No significant change 1 = 1 grade progressions (worsening) 2 = 2 grade progressions (worsening) -1 = 1 grade regression (improvement) -2 = 2 grade regression (improvement)
ECHG_INF	Year 1-4 change in STL_INF	ST Elevation change inferior	
ECHG_LAT	Year 1-4 change in STL_LAT	ST Elevation change lateral	
FLAGECG	Flag set to 1 when participant has ECG data	Flag set to 1 when ppt has ECG data	1= Has ECG data

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
HABCID	Health ABC Enrollment ID #	Enrollment ID	
HR	Heart rate	Heart rate	beats per minute
LONGPR	Long P-R interval	Indicator variable for P-R interval > 220 msec	See Y4Calc
LONGQRS	Long QRS duration	QRS duration longer than 120 msec	See Y4Calc
LONGQT	Long QT duration	QT duration longer than 460 msec	See Y4Calc
LVH	Left ventricular hypertrophy code	Left ventricular hypertrophy code	1 =Definite LVH by voltage (R amplitude >26 mm in either V5 or V6, or R amplitude >20.0 mm in any of leads I, II, or III, aVF, or R amplitude >12.0 mm in lead aVL measured only on second to last complete normal beat) 3 = Probable LVH by voltage (optional code when 1 is not present: R amplitude >15.0 mm but ≤ 20.0 mm in lead I, or R amplitude in V5 or V6, plus S amplitude in V1 > 35.0 mm)
PR	PR interval	PR interval	msec Note: blank if not assessed
Q_ANT	Q and QS patterns anterior site (leads VI, V2, V3, V4, V5): To qualify as a Q-wave, the deflection should be at least 1.0 mv (amplitude).	Q code anterior	Major: 11 = Q/R amplitude ratio ≥ 1/3, plus Q duration ≥ 0.03 sec in any of leads V2, V3, V4, V5 12 = Q duration ≥ 0.04 sec in any of leads V1, V2, V3, V4, V5 16 = QS pattern when initial R-wave is presented in adjacent lead to the right on the chest, in any of leads V2, V3, V4, V5, V6 17 = QS pattern in all of leads V1-V4, or V1-V5 21 = Q/R amplitude ratio ≥ 1/3 plus Q duration ≥ 0.02 sec and 0.03 sec, in any of leads V2, V3, V4, V5 22 = Q duration ≥ 0.03 sec and < 0.04 sec in any of leads V2, V3, V4, V5 27 = QS pattern in all of leads V1, V2, and V3 Minor: 28 = Initial R amplitude decreasing to 2 mm or less in every beat between any of the leads V2 and V3, V3 and V4, or V4 and V5 (All beats in the lead immediately to the right of the chest must have an initial R amplitude > 2 mm.) 31 = Q/R amplitude ratio ≥ 1/5 and < 1/3, plus Q duration ≥ 0.02 sec and < 0.03 sec in any of leads V2, V3, V4, V5 32 = QS pattern in lead V1 and V2

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
Q_INF	Q and QS patterns inferior site (leads II, III, aVF): To qualify as a Q-wave, the deflection should be at least 1.0 mv (amplitude).	Q code inferior	<p>Major:</p> <p>11 = Q/R amplitude ratio \geq 1/3, plus Q duration \geq 0.03 sec in lead II</p> <p>12 = Q duration \geq 0.04 sec in lead II</p> <p>14 = Q duration \geq 0.05 sec in lead III, plus Q wave amplitude \geq 1 mm in the majority of beats in aVF</p> <p>15 = Q duration \geq 0.05 sec in aVF</p> <p>21 = Q/R amplitude ratio \geq 1/3, plus Q duration \geq 0.02 sec and $<$ 0.03 sec in lead II</p> <p>22 = Q duration \geq 0.03 sec and $<$ 0.04 sec in lead II</p> <p>23 = QS pattern in lead II</p> <p>24 = Q duration \geq 0.04 sec and $<$ 0.05 sec in lead III, plus Q-wave \geq 1.0 mm amplitude in the majority of beats in aVF</p> <p>25 = Q duration \geq 0.04 sec and $<$ 0.05 sec in lead aVF</p> <p>Minor:</p> <p>26 = Q amplitude \geq 5.0 mm in leads III or aVF</p> <p>31 = Q/R amplitude ratio \geq 1/5 and $<$ 1/3, plus Q duration \geq .02 sec and $<$ 0.03 sec in lead II</p> <p>34 = Q duration \geq 0.03 sec and $<$ 0.04 sec in lead III, plus a Q wave \geq 1.0 mm amplitude in the majority of beats in aVF</p> <p>35 = Q duration \geq 0.03 sec and $<$ 0.04 sec in lead aVF</p> <p>36 = QS pattern in each of leads III and aVF</p>
Q_LAT		Q code lateral	<p>Major:</p> <p>11 = Q/R amplitude ratio \geq 1/3, plus Q duration \geq 0.03 sec in lead I or V6</p> <p>12 = Q duration \geq 0.04 sec in lead I or V6</p> <p>13 = Q duration \geq 0.04 sec, plus R amplitude \geq 3 mm in lead aVL</p> <p>21 = Q/R amplitude ratio \geq 1/3, plus Q duration \geq 0.02 sec and 0.03 sec in lead I or V6</p> <p>22 = Q duration \geq 0.03 sec and $<$ 0.04 sec in lead I or V6</p> <p>23 = QS pattern in lead I</p> <p>Minor:</p> <p>28 = initial R amplitude decreasing to 2 mm or less in every beat between V5 and V6 (all beats in V5 must have an initial R amplitude $>$ 2 mm)</p> <p>31 = Q/R amplitude ratio \geq 1/5 and $<$ 1/3, plus Q duration \geq .02 sec and $<$ 0.03 sec in I or V6</p> <p>33 = Q duration \geq 0.03 sec and $<$ 0.04 sec, plus R amplitude \geq 3 mm in aVL</p>

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
Q_POST	Q and QS patterns posterior site	Q code posterior	Major: 1 = R duration in V1 \geq 40 and (R height V1/S depth V1) > 1 Minor: 2 = R duration in V2 \geq 40 and (R height V2/S depth V2) >1
QCHG_ANT	Year 1-4 change in Q_ANT	Q change anterior	0 = No significant change 1 = 1 grade progressions (worsening) 2 = 2 grade progressions (worsening) -1 = 1 grade regression (improvement) -2 = 2 grade regression (improvement)
QCHG_INF	Year 1-4 change in Q_INF	Q change inferior	
QCHG_LAT	Year 1-4 change in Q_LAT	Q change lateral	
QCHG_POS	Year 1-4 change in Q_POS	Q change posterior	
QMI	Q-wave myocardial infarction	Presence of major Q-wave abnormality or minor Q-wave abnormality in the presence of a major ST or T-wave abnormality	See Y4Calc
QRS	QRS duration	QRS duration	msec
QT	QT interval	QT interval	msec
QTC	QT interval corrected for heart rate	QT interval corrected for heart rate	msec
QWVAB	Q-wave myocardial infarction	Major Q or QS abnormality for anterior, posterior, lateral or inferior	See Y4Calc
SITE*	Field Center	Field Center	1 = Memphis; 2 = Pittsburgh

* Must link to Health ABC participant history file (PH) to add this variable.

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
STD_ANT	ST junction (J) and segment depression anterior site (leads V1, V2, V3, V4, V5)	ST depression anterior	<p>Major:</p> <p>11 = STJ depression ≥ 2.0 mm and ST segment horizontal or downward sloping in any of leads V1, V2, V3, V4, V5</p> <p>12 = STJ depression ≥ 1.0 mm but < 2.0 mm and ST segment horizontal or downward sloping in any of leads V1, V2, V3, V4, V5</p> <p>13 = STJ depression ≥ 3.0 mm and ST segment horizontal downward sloping</p> <p>14 = STJ depression ≥ 4.0 mm and ST segment horizontal or downward sloping</p> <p>1X = STJ depression $\geq X.0$ mm and ST segment horizontal or downward sloping. X must be ≥ 5.0 mm</p> <p>Minor:</p> <p>2 = STJ depression ≥ 0.5 mm and < 1.0 mm and ST segment horizontal or downward sloping in any of leads V1, V2, V3, V4, V5</p> <p>3 = No STJ depression as much as 0.5 mm, but ST segment downward sloping and segment or T-wave nadir ≥ 0.5 mm below P-R baseline, in any of leads V2, V3, V4, V5</p> <p>4 = STJ depression ≥ 1.0 mm and ST segment upward sloping or U-shaped, in any of leads V1, V2, V3, V4, V5</p>
STD_INF	ST junction (J) and segment depression anterior site (leads II, III, aVF)	ST depression inferior	<p>Major:</p> <p>11 = STJ depression ≥ 2.0 mm and ST segment horizontal or downward sloping in lead II or aVF</p> <p>12 = STJ depression ≥ 1.0 mm but < 2.0 mm and ST segment horizontal or downward sloping in lead II or aVF</p> <p>Minor:</p> <p>2 = STJ depression ≥ 0.5 mm and < 1.0 mm and ST segment horizontal or downward sloping in lead II or aVF</p> <p>3 = No STJ depression as much as 0.5 mm, but ST segment downward sloping and segment or T-wave nadir ≥ 0.5 mm below P-R baseline in lead II</p> <p>4 = STJ depression ≥ 1.0 mm and ST segment upward sloping, or U-shaped, in lead II</p>
STD_LAT (cont next page)	ST junction (J) and segment depression lateral site (leads I, aVL, V6)	ST depression lateral	<p>Major:</p> <p>11 = STJ depression ≥ 2.0 mm and ST segment horizontal or downward sloping in any of leads I, aVL, or V6</p> <p>12 = STJ depression ≥ 1.0 mm and < 2.0 mm and ST segment horizontal or downward sloping in any of leads I, aVL, or V6</p> <p>(cont next page)</p>

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
STD_LAT (cont from previous page)			(cont from previous page) Minor: 2 = STJ depression ≥ 0.5 mm and < 1.0 mm and ST segment horizontal or downward sloping in any of leads II or aVF 3 = No STJ depression as much as 0.5 mm, but ST segment downward sloping and segment or T-wave nadir ≥ 0.5 mm and below P-R baseline, in any of leads I, aVL, or V6 4 = STJ depression ≥ 1.0 mm and ST segment upward sloping or U-shaped, in any of leads I, aVL, or V6
STL_ANT	ST segment elevation anterior site (leads V1, V2, V3, V4, V5)	ST segment elevation anterior site	Major: 2 = STJ segment elevation ≥ 1.0 mm in lead V5 or ST segment elevation ≥ 2.0 mm in leads V1, V2, V3, V4 21 = STJ segment elevation ≥ 1.5 mm in lead V5 or ST segment elevation ≥ 2.5 mm in leads V1, V2, V3, V4 2X = STJ segment elevation $\geq X.0$ mm in leads V1, V2, V3, V4. X must be ≥ 2.0 mm.
STL_INF	ST segment elevation inferior site (leads II, III, aVF)	ST segment elevation inferior site	Major: 2 = STJ segment elevation ≥ 1.0 mm in leads II, III, aVF 21 = STJ segment elevation ≥ 1.5 mm in leads II, III, aVF 2X = STJ segment elevation $\geq X.0$ mm in leads II, III, aVF. X must be ≥ 2.0 mm.
STL_LAT	ST segment elevation lateral site (leads I, aVL, V6)	ST segment elevation lateral site	Major: 2 = STJ segment elevation ≥ 1.0 mm in leads I, aVL, V6 21 = STJ segment elevation ≥ 1.5 mm in leads II, aVL, V6 2X = STJ segment elevation $\geq X.0$ mm in leads I, aVL, V6. X must be ≥ 1.5 mm.
STSEGDEP	ST segment depression	ST segment depression excluding Q-wave myocardial infarction	See Y4Calc
STSEGELV	ST segment elevation	ST segment elevation excluding Q-wave myocardial infarction	See Y4Calc
STWVAB	ST or T wave abnormality	Major ST or T wave abnormality for anterior, posterior, lateral or inferior	See Y4Calc

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
SUPPCODE	Is a suppression code present that invalidates the comparison?	Suppression code invalidates comparison?	0 = No 78 = No, right bundle branch block w/left axis dev. 721 = No, right bundle branch block 1 = Yes, Missing ECG 2 = Yes, lead reject all 31 = Yes, left ventricular hypertrophy 68 = Yes, pacemaker 641 = Yes, Wolff-Parkinson=White 711 = Yes, Left bundle branch block 821 = Yes, ventricular fibrillation 841 = Yes, super-ventricular rhythm 964 = Yes, torso ECG 965 = Yes, torso ECG w/invalid calibration 981 = Yes, technical prob. interferes w/coding 984 = Yes, photocopy, tech probs interferes w/coding
T_ANT	T-wave items anterior site (leads V2, V3, V4, V5)	T-wave items anterior site	Major: 1 = T amplitude negative 5.0 mm or more in any of leads V2, V3, V4, V5 5 = T amplitude negative (flat) diphasic (negative-positive or positive-negative type) with negative phase at least 1.0 mm in at least three leads V2, V3, V4, V5 6 = T amplitude negative 5.0 mm or more in at least three leads V2, V3, V4, and V5 Minor: 2 = T amplitude negative (flat) or diphasic (positive-negative or negative-positive type) with negative phase at least 1.0 mm but not as deep as 5.0 mm in any leads V2, V3, V4, V5 3 = T amplitude zero (flat) or negative, or diphasic (negative-positive type only) with less than 1.0 mm negative phase, in any of leads V3, V4, V5 4 = T amplitude positive and T/R amplitude ratio < 1/20 in any of leads V3, V4, V5; R wave amplitude must be ≥ 10.0 mm
T_INF	T-wave items inferior site (leads II, III, aVF)	T-wave items inferior site	Major: 1 = T amplitude negative 5.0 mm or more in lead II or in lead aVL when QRS is mainly upright 2 = T amplitude negative or diphasic (positive-negative or negative-positive type) with negative phase at least 1.0 mm but not as deep as 5.0 mm in lead II or in lead aVF when QRS is mainly upright Minor: 3 = T amplitude zero (flat) or negative, or diphasic (negative-positive type only) with less than 1.0 mm negative phase in lead II; not coded in lead aVF 4 = T amplitude positive and T/R amplitude ratio < 1/20 in lead II; R wave amplitude must be ≥ 10.0 mm

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
T_LAT	T-wave items lateral site (leads I, aVL, V6)	T-wave items lateral site	<p>Major:</p> <p>1 = T amplitude negative 5.0 mm or more in either of leads I, V6, or in lead aVL when R amplitude is \geq 5.0 mm</p> <p>2 = T amplitude negative or diphasic (positive-negative or negative-positive type) with negative phase at least 1.0 mm but not as deep as 5.0 mm in leads I or V6, or in leads aVL when R amplitude is \geq 5.0 mm</p> <p>Minor:</p> <p>3 = T amplitude zero (flat) or negative, or diphasic (negative-positive type only) with less than 1.0 mm negative phase in lead I or V6, or in aVL when R amplitude is $>$ 5.0 mm</p> <p>4 = T amplitude positive and T/R amplitude ratio $<$ 1/20 in any of leads I, aVL, V6: R wave amplitude must be \geq 10.0 mm</p>
TCHG_ANT	Year 1-4 change in T_ANT	T wave change anterior	<p>0 = No significant change</p> <p>1 = 1 grade progressions (worsening)</p> <p>2 = 2 grade progressions (worsening)</p> <p>-1 = 1 grade regression (improvement)</p> <p>-2 = 2 grade regression (improvement)</p>
TCHG_INF	Year 1-4 change in T_INF	T wave change inferior	<p>0 = No significant change</p> <p>1 = 1 grade progressions (worsening)</p> <p>2 = 2 grade progressions (worsening)</p> <p>-1 = 1 grade regression (improvement)</p> <p>-2 = 2 grade regression (improvement)</p>
TCHG_LAT	Year 1-4 change in T_LAT	T wave change lateral	
TECHFLAG	Flag for existence of technical problem interfering with coding	If TECH_PRB=981 or 984, then flag=1. All others, flag=0	See Y4Calc
TECH_PRB	ECG technical problem	ECG technical problem	<p>981 = technical problems which interfere with coding</p> <p>982 = technical problems which do not interfere with coding</p> <p>983 = photocopy. No technical problems which interfere with coding</p> <p>984 = photocopy. Technical problems which interfere with coding.</p>
TWVITMS	T-wave items	T-wave items excluding Q-wave myocardial infarction	See Y4Calc
VDEF COD	Ventricular conduction defect code	Combines VENT_DEF=12, 22, 3, 5, and 6; keeps 11, 21, 4, 7 and 8 separate	See Y4Calc

RESTING ECG DATA

Variable Name	Variable Description	Variable Label	Value Label
VENT_DEF	Ventricular conduction defect	Ventricular conduction defect code	11 = complete left bundle branch block (LBBB) 12 = intermittent left bundle branch block 21 = complete right bundle branch block (RBBB) 22 = intermittent right bundle branch block 3 = incomplete right bundle branch block. QRS duration < .12 sec in each of leads I, II, III, aVL, and R'>R in either of leads V1, V2. 4 = intraventricular block 5 = R-R' pattern in either of leads V1, V2 with R' amplitude \leq R 6 = incomplete left bundle branch block 7 = left anterior hemiblock (LAH) 8 = combination of 7 and 21 or 7 and 22

**HEALTH ABC READING CENTER VARIABLE LIST
ISOKINETIC STRENGTH (KINCOM)**

Variable Name	Variable Description	Variable Label	Value Label
FLAGKC	Flag set to 1 when participant has KinCom data	Flag set to 1 when ppt has KinCom data	
HABCID	Health ABC Enrollment ID #	Enrollment ID	
KCARM	Dynamometer arm length	Dyno arm length	cm
KCEDATE	Exam date: KinCom	Exam date: KinCom	MM/DD/YY
KCFARM	Dynamometer arm - final position	Dyno arm: final angle	°
KCFMAX	Maximum force (moving from 80 to 40 degrees) from averaged curves	Avg maximum force (N): 80-40deg	N
KCFMEAN	Average force (moving from 80 to 40 degrees) from averaged curves	Average force (N): 80-40deg	N
KCFMIN	Minimum force from averaged curves (in 80 to 40 degree range)	Avg minimum force (N): 80-40 deg	N
KCFSTD	Standard deviation (force) from averaged curves (in 80 to 40 degree range)	Standard deviation (force) (N)	
KCSARM	Dynamometer arm: initial angle	Dyno arm: initial angle	°
KCSTID	Staff ID #: KinCom	Staff ID (KinCom)	
KCTMAX	Maximum torque (moving from 80 to 40 degrees) from averaged curves	Avg max torque (Nm): 80-40deg	Nm
KCTMEAN	Average torque (moving from 80 to 40 degrees) from averaged curves	Average torque (Nm): 80-40deg	Nm
KCTMIN	Minimum torque from averaged curves (in 80 to 40 degree range)	Average minimum torque (Nm): 80-40 deg	Nm
KCTSTD	Standard deviation (torque) from averaged curves (in 80 to 40 degree range)	Standard deviation (torque) (Nm)	Nm
SITE*	Field center	Field Center	1=Memphis 2=Pittsburgh
TRIAL	Number of trials	# of trials	

* Must link to Health ABC participant history file (PH) to add this variable.

**PERIPHERAL MOTOR NEURON CONDUCTION DATA
HEALTH ABC READING CENTER VARIABLE LIST**

Variable Name	Variable Description	Label	Units
ADIST	Distance from ankle cathode to recording electrode	Ankle to recording electrode distance	mm
FHDIST	Distance from ankle cathode to fibular head cathode	Fibular head to ankle electrode distance	mm
FLAGPMNC	Flag set to 1 when participant has PMNC data	Flag set to 1 when ppt has PMNC data	1=Has PMNC data
HABCID	Health ABC Enrollment ID#	HABC Enrollment ID	
PFDIST	Distance from ankle cathode to Popliteal fossa cathode	Pop foss to ankle electrode distance	mm
Y4AAMP	Ankle amplitude	Y4 Ankle Amplitude (mV)	mV
Y4AAMPFX	Flag for Y4 ankle amplitude (low or normal range)	Y4 Ankle Amplitude Flag (0: <1.0, Low; 1: >=1.0, Normal)	0 = <1 mV (Low) 1 = > 1 (Normal)
Y4FHAMP	Fibular head amplitude	Y4 Fibular Head Amplitude (mV)	mV
Y4FHAMPFX	Flag for Y4 fibular head amplitude (low or normal range)	Y4 Fibular Head Amplitude Flag (0: <1.0, Low; 1: >=1.0, Normal)	0 = <1 mV (Low) 1 = > 1 (Normal)
Y4FHCV	Conduction velocity between fibular head and ankle	Y4 Fibular Head Conduction Velocity (m/s)	m/s
Y4FHCVFX	Flag for Y4 fibular head conduction velocity (low, normal or high)	Y4 Fibular Head Conduction Velocity Flag (0: <20.0, Low; 1: >=20.0 and <=70.0, Normal; 2: >70.0, High)	0 = <20 m/s (Low) 1 = 20-70 m/s (Normal) 2 = >70 m/s (High)
Y4FHNC_ALL	Flag for valid data for both amplitude and velocity for Y4 fibular head nerve conduction data	Y4 Fibular Head Nerve Conduction Valid Data for both Amplitude and Velocity (0=No; 1=Yes)	1=Yes
Y4PFAMP	Popliteal fossa amplitude	Y4 Popliteal Fossa Amplitude (mV)	mV

PMNC DATA

Variable Name	Variable Description	Label	Units
Y4PFAMPFX	Flag for Y4 popliteal fossa amplitude (low or normal range)	Y4 Popliteal Fossa Amplitude Flag (0: <1.0, Low; 1: >=1.0, Normal)	0 = <1 mV (Low) 1 = > 1 (Normal)
Y4PFCV	Conduction velocity between Popliteal fossa and ankle	Y4 Popliteal Fossa Conduction Velocity (m/s)	m/s
Y4PFCVFX	Flag for Y4 popliteal fossa conduction velocity (low, normal or high)	Y4 Popliteal Fossa Conduction Velocity Flag (0: <20.0, Low; 1: >=20.0 and <=70.0, Normal; 2: >70.0, High)	0 = <20 m/s (Low) 1 = 20-70 m/s (Normal) 2 = >70 m/s (High)
Y4PFNC_ALL	Flag for valid data for both amplitude and velocity for Y4 popliteal fossa nerve conduction data	Y4 Popliteal Fossa Nerve Conduction Valid Data for both Amplitude and Velocity (0=No; 1=Yes)	1=Yes

**HEALTH ABC READING CENTER VARIABLE LIST
VIBRATION PERCEPTION THRESHOLD**

Variable Name Variable Description Variable Label Value Label

D2VPTOE	Toe tested according to form data	Y4CVW: VPT: Q24: Which great toe...tested?	1=Right 2=Left
FLAGNRVB	Flag set to 1 when participant has nerve vibration data	Flag set to 1 when ppt has neuro vibrati data	1=Has nerve vibration data
VIBADAP	Baseline value for vibration test	Vibration baseline value for test	microns
VIBAVG	Average perceived stimulus	Average vibration perceived	microns
HABCID	Health ABC Enrollment ID#	Health ABC Enrollment ID#	
VIBSTFID	Staff ID for vibration testing	Vibration Staff ID	
VIBSRC	Source of the data	Vibration data source 1=Elect 2=Form	1=Electronic 2=Form
VBSTART1	Stimulus felt, first measurement	Vibration stimulus felt, meas 1	microns
VBSTART2	Stimulus felt, second measurement	Vibration stimulus felt, meas 2	microns
VBSTART3	Stimulus felt, third measurement	Vibration stimulus felt, meas 3	microns
VBSTART4	Stimulus felt, fourth measurement	Vibration stimulus felt, meas4	microns
VBSTART5	Stimulus felt, fifth measurement	Vibration stimulus felt, meas 5	microns
VIBTIME	Time of day, vibration measurement (24-hour)	Vibration time (24-hr)	
VIBTOE	Which toe tested for vibration threshold	Vibration toe tested	
VIBVAR	Variance of vibration measurements	Vibration variance	microns ²