CAROTID ULTRASOUND SCANS
Visit 12 with Visit 13 Updates and Visit 15

CODEBOOK

ARCHIVED DATASET 2019

PAGE INTENTIONALLY BLANK

DOCUMENTATION FOR THE SWAN VISIT 12 WITH VISIT 13 UPDATES AND VISIT 15 CAROTID SCAN DATA

Changes from the prior release of SWAN Carotid Scan dataset (niacarotid2018): V15 data has been added.

***Who is included in the frozen dataset*:**
SWAN participants who were active as of follow-up visits 12,13 or 15 and who completed a visit 12, visit 13 or visit 15 carotid assessment at the six participating SWAN study sites are included in the frozen dataset. Sixteen observations (0.65%) out of the 2457 women completing visit 12, 13 or 15 carotid artery assessments were dropped after data cleaning, resulting in 2441 women in the frozen data. Of these women, 2435 had complete plaque yes/no (PLAQUE) data, 2410 had complete plaque index (**PI**) data, 2423 women had some carotid IMT data and 2342 of them had complete mean average CCA IMT (**MAVG_CCA**) data, the primary variable generated from reading of the carotid scan images.

SWAN participants from the Chicago, Pittsburgh and New Jersey sites who did *not* receive a carotid scan during Visit 12 were invited to undergo a carotid scan for Visit 13. All eligible participants at the four participating sites at Visit 15 were brought in for additional carotid scan, except for 43 participants who were brought in at Visit 15 for a carotid scan for the first time.

Ultrasound technology was used to obtain the carotid measures in the following datasets. This frozen dataset represents a merging of 2 raw datasets:

1) **IMT worksheet** raw data**:** Centrally trained and certified site sonographers acquired the carotid artery scan images, assessed presence of plaque and recorded results on the URL Carotid IMT Worksheet form. Carotid plaque data from this form were then entered by each site.
2) **IMT read** raw data**:** Carotid scan images were then data streamed to the SWAN Ultrasound Reading Center (University of Pittsburgh URL) for centralized reading using semi-automated edge detection software. Common carotid artery (CCA), intima-media thickness (IMT), and adventitial diameter (AD) data were obtained from these readings and were processed and provided by the Ultrasound Reading Center to SWAN data management.

Breakdown of study participants with carotid scan data by site and visit

	Sites (site number)						All Sites
	Detroit (11)	Boston (12)	Chicago (13)	Oakland (14)	Newark (16)	Pittsburgh (17)	
Visit 12							
Completed Carotid visit	362	206	246	341	139	272	1566
Included in frozen dataset	350	206	246	341	139	270	1552
Complete CCA IMT data	327	200	241	337	135	267	1507
Complete PI data	350	205	246	341	139	269	1550
Visit 13							
Completed Carotid visit	--------	--------	23	--------	8	23	54
Included in frozen dataset	--------	--------	23	--------	8	23	54
Complete CCA IMT data	--------	--------	20	--------	8	23	51
Complete PI data	--------	--------	23	--------	8	23	54

Visit 15							
	Sites (site number)						All Sites
	Detroit (11)	Boston (12)	Chicago (13)	Oakland (14)	Newark (16)	Pittsburgh (17)	
Completed Carotid visit	266	--------	185	--------	124	262	837
Included in frozen dataset	264	--------	183	--------	124	262	833
Complete CCA IMT data	262	--------	184	--------	124	261	831
Complete PI data	265	--------	184	--------	124	262	835
Visit 12+13 and Visit 15 in frozen dataset	615	206	453	341	271	555	2441

***Data cleaning strategy**:*
The strategy for cleaning URL carotid ultrasound data occurred in three stages. In stage one (before images were read), carotid images were reviewed and scored for overall quality assessment by the URL.

In stage two (after images were read), carotid ultrasound data were reviewed quarterly during data collection. During this stage, potential outliers for the variables average CCA IMT (**mavg_cca**) and average adventitial diameter (**adavg**) were flagged and queried by the URL. Potential outliers were identified as observations with values outside mean ± 2SD of site specific quarterly data. Random data checks were also performed throughout data collection and anomalies were queried. The query involved reviewing the URL Carotid IMT Worksheets and log sheets completed by site sonographers as well as the carotid images obtained at scan time. These quarterly data checks were compiled and finalized once data collection was completed for the visit at all sites. Please note that at this stage plaque index (**PI**) was not used for flagging observations to be checked but was reviewed for specific observations when mavg_cca or adavg values were flagged.

The third stage of data cleaning involved the review of the preliminary carotid ultrasound dataset as a whole at the end of visit 12, visit 13 and visit 15. For this review, the strategy was to flag observations with values falling outside the soft range listed in Table 2 for the following variables: **mavg_cca, adavg** and **pi**. In addition, values outside the range mean ± 2SD for continuous measures were also flagged. The observations flagged were then cross-checked with the previously compiled list of quarterly data checks. Any observations not previously queried were checked at this time.

Table 2: Soft range for data cleaning and checking

Variable	Soft range*	
	Visit 12 and 13	Visit 15
mavg_cca	0.55-1.00	0.55-1.25
adavg	5.5-9.0	5.5-9.5
pi	0 - 8	0 - 8

* Values outside the soft range were flagged for data checking/cleaning

***Summary of Data Check/Cleaning**:*
A brief summary of the data checks performed and the findings are as follows:

- **Fourteen observations were dropped:** Initially, 33 observations were dropped because the images were considered unreadable or of very poor quality for IMT and AD measurement per the Reading Center protocol (n=31), or images could not be obtained due to technical machine malfunction (n=2). Later on, 17 of these 33 observations were reincorporated into the frozen dataset because the scans were deemed adequate for the assessment of presence of plaque, resulting in N=2435 observations with plaque data and N=2423 observations with any IMT data.

- **Fifty Three observations were modified in raw data**: For some observations that were flagged, re-reading of the images were recommended and resulted in data of higher quality. For these, re-read data replaced the original read data as indicated in the data check table.
- **Observations flagged for Analyses**: For flagged observations, please follow recommendations indicated by flag codes. It is recommended that analysts acknowledge within the analyses that these data may be inaccurate. An assessment of the quality of these observations could be done in several ways, and options include running sensitivity analyses (with and without these data flagged), dropping the observations, or acknowledging in methods these specific measurement errors. The final decision is up to the investigator/analyst. Flag variables in the frozen dataset were only created for primary data collected at all sites, i.e., MAVG_CCA and ADAVG.

Changes Made to Data to Ensure Standard Coding:

- Range checks, skip pattern checks and other edits were run on the data during the data entry process. Errors were resolved with the sites or Reading Center during this process. Therefore, few changes were made to the data during the creation of the SAS dataset.
- A site variable (**SITE**) was added to the dataset to indicate whether the participant was from
- A SWAN visit indicator variable (**VISIT)** was added to the dataset to indicate that the carotid data were collected at visit 12 (**VISIT**=12), visit 13 (**VISIT**=13) or visit 15 (**VISIT**=15).
- Coding of Missing Values: The original missing codes (.B, .C, -9) from the raw data were changed to the SAS missing values (".")

***Created* Flag Variables**:
The following flag variables were created based on review of the data.

Name	Meaning	Code
F LGCCA	If yes, observations should be **used with caution** in analyses of **mavg_cca** because the data may not be valid; specified in data check table.	0 = No, 1 = Yes
FLGAD	If yes, observations should be **used with caution** in analyses of **adavg** because the data may not be valid; specified in data check table.	0 = No 1 = Yes
CALCCA	If yes, observations are missing data for **mavg_cca**. Analyst may estimate the value of **mavg_cca** with caution by using available data and following created variable equation for **mavg_cca.**	0 = No 1 = Yes
CALAD	If yes, then observations are missing data for **adavg**. Analyst may estimate value of **adavg** with caution by using available data and following created variable equation for **adavg.**	0 = No 1 = Yes
FLAG_pi	If yes, the observation should be **used with caution** in analyses using plaque index. Options are to not use these observations for plaque index analyses or alternatively to use the imputed variables, **pi_min, pir_min and pil_min** for analysis.	0 = No 1 = Yes

Missing Plaque index:
The following observations were flagged by the Reading Center for having a missing plaque index (**pi**). For many of these observations, the site sonographer failed to enter "0" for plaque grade, which resulted in missing data for **pi**. The verification process performed by the Reading Center noted that these participants did not have plaque and therefore **the variables pi** and **plaque** were set to "0".

Observations with PLAQUE data but no carotid IMT data in frozen dataset:
Seventeen observations were brought back into the frozen dataset because although their carotid images were not adequate for the inclusion of measures of IMT or AD they were adequate for the assessment of absence/presence of plaque.

URL CAROTID IMT WORKSHEET WITH VARIABLE NAMES

Note: Two different versions of this form were used to collect the plaque data. Two versions of the form were implemented - Oct 13th 2009 and Mar 9th 2010. If the older form (Oct 13th 2009) was used, then the variable PVISUALR and PVISUALL will be missing. If the new form (Mar 9th 2010) was used, then data for the variable PVISUALR and PVISUALL was collected. In this case, if a site sonographer did not see any plaque, these variables were coded as "0" and data for the corresponding questions #3, #4 and #5 in section A of the form were skipped and coded as ".B", the SAS missing codes for "N/A". The creation of plaque variables was dependent on form version.

URL Carotid IMT Worksheet

Ultrasound Research Lab

URL Carotid IMT Worksheet	
Study: #STUDY	URL Visit #: VISIT
Study ID: ARCHID~	Enter:
URL ID: #LABID	Verify:
Tech ID: #TECH **Test Date**: SCANDAY†	**CD #:** #CD
Note: Technologist is responsible for bold faced data above.	***Repro only:*** **Test Seq:** #SCAN

Plaque Reading Software Used? Yes ( ) No ( ) #PRSOFT

RIGHT CAROTID ARTERY

A. Plaque Index – Right

	1-Prox CCA	2-Distal CCA	3-Bulb	ICA	ECA
1. Segment adequately visualized	Y ( ) N ( ) #VIS1R	Y ( ) N ( ) #VIS2R	Y ( ) N ( ) #VIS3R	Y ( ) N ( ) #VIS4R	Y ( ) N ( ) #VIS6R
2. Any plaques visualized	Y ( ) proceed to 3		N ( ) proceed to 6	#PVISUALR	
3. No. of Lesions	#LES1R	#LES2R	#LES3R	#LESIR	#LESER
4. Plaque Grade* (0, 1, 2, 3)	#GRA1R	#GRA2R	#GRA3R	#GRAIR	#GRAER
5. Calcified Plaque	Y ( ) N( ) #CAP1R	Y ( ) N( ) #CAP2R	Y ( ) N( ) #CAP3R	Y ( ) N( ) #CAPIR	Y ( ) N( ) #CAPER

Velocities done if Plaque Grade is ≥ 3 in any segment; please complete URL Carotid Doppler Velocity Worksheet

~ A randomly generated ID will be provided that is different from the original ID.
† This date is given in days since the initial baseline interview, which is day zero.

 # Variable Excluded from Public Use Data File

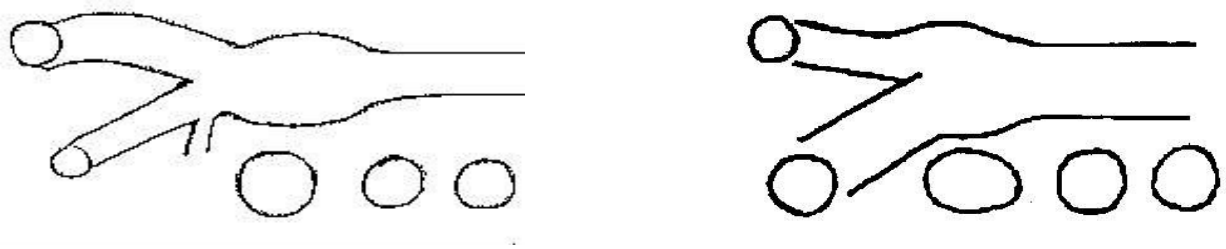
6. Comments:

Technically Difficult Study ❑ **YES*** ↓ ❑ **NO** #TDSR		**Machine Failure** ❑ **YES*** ↓ ❑ **NO** #MECHR
❑ Tortuous vessel #TVR	❑ Deep Vessels #DVR	Reason for machine failure: #REASONR
❑ Participant movement #PMR	❑ Morbidly obese #MOR	

***Comments required:** _____#COMMENTR_____

B. Plaque and Wall thickening – Right

ICA **ECA**



Ultrasound Research Lab

URL ID: #LABID **Test Date:** SCANDAY†

LEFT CAROTID ARTERY

A. Plaque Index – Left

	1-Prox CCA	2-Distal CCA	3-Bulb	ICA	ECA
1. Segment adequately visualized	Y ( ) N ( ) #VIS1L	Y ( ) N ( ) #VIS2L	Y ( ) N ( ) #VIS3L	Y ( ) N ( ) #VIS4L	Y ( ) N ( ) #VIS6L
2. Any plaques visualized	Y ( ) proceed to 3	N ( ) proceed to 6	#PVISUALL		
3. No. of Lesions	#LES1L	#LES2L	#LES3L	#LESIL	#LESEL
4. Plaque Grade* (0, 1, 2, 3)	#GRA1L	#GRA2L	#GRA3L	#GRAIL	#GRAEL
5. Calcified Plaque	#CAP1L Y ( ) N( )	#CAP2L Y ( ) N( )	#CAP3L Y ( ) N( )	#CAPIL Y ( ) N( )	# Y ( ) N( )

Velocities done if Plaque Grade is ≥ 3 in any segment; please complete URL Carotid Doppler Velocity Worksheet

† This date is given in days since the initial baseline interview, which is day zero.
 # Variable Excluded from Public Use Data File

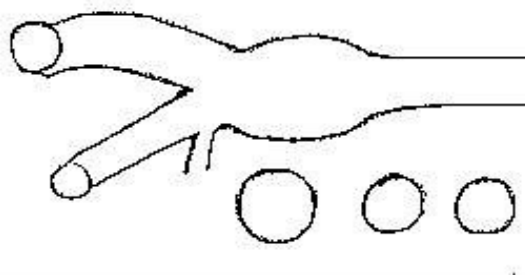
6. Comments:

Technically Difficult Study ❑ **YES*** ↓ ❑ **NO** **#TDSL**		**Machine Failure** ❑ **YES*** ↓ ❑ **NO** **#MECHL**
❑ Tortuous vessel **#TVL**	❑ Deep Vessels **#DVL**	Reason for machine failure: **#REASONL**
❑ Participant movement **#PML**	❑ Morbidly obese **#MOL**	

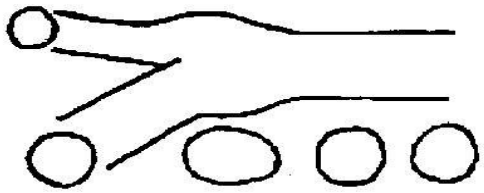
***Comments required: _____ #COMMENTL_____**

B. <u>Plaque and Wall thickening</u> – Left

ICA



ECA



CREATED VARIABLES

Main IMT and AD Variables for Analyses:
The following IMT and AD variables were created from data yielded from reading of the carotid images using semi- automated edge detection software. Equations for created variables are listed in the next section. **The units for all these variables are millimeters (mm)**.

Variables	Label
mavg_cca	Mean IMT of Average CCA
mavgr_cca	Mean IMT of Average CCA –Right
mavgl_cca	Mean IMT of Average CCA –Left
mmaxr_cca	Max IMT of Average CCA-Right
mmaxl_cca	Max IMT of Average CCA-Left
mmax_cca	Max IMT of Average CCA
mminr_cca	Minimum IMT of Average CCA-Right
mminl_cca	Minimum IMT of Average CCA-Left
mmin_cca	Minimum IMT of Average CCA
adavg	Mean average CCA adventitial diameter
adavgr	Mean average CCA adventitial diameter-right
adavgl	Mean average CCA adventitial diameter-left
admin	Mean minimum CCA adventitial diameter
admax	Mean maximum CCA adventitial diameter
lumen	Lumen Diameter
lumenr	Lumen Diameter – Right
lumenl	Lumen Diameter - Left

Carotid intima-media thickness or IMT represents the thickness of the innermost two layers of the carotid artery walls and inter-adventitial diameter or AD represents the diameter of the common carotid artery. This diameter measurement includes the lumen and the intimal and medial layers of the CCA.

Main Plaque Variables for Analyses:
The following plaque variables were created from variables derived from URL Carotid IMT Worksheet.

Variables	Label
pir	Right plaque index
pil	Left plaque index
pi	Aggregate of left and right plaque index
Plaque	Absence or presence of plaque

Plaque Index:
The plaque index is the summation of the plaque grades across all carotid segments visualized for both left and right carotid arteries. Plaque grade is an estimate of the extent of focal plaque in the carotid artery segments visualized. Each carotid segment (CCA, Bulb, ICA and ECA) is assessed and scored individually using the criteria that follows.

Estimated extent of focal plaque	**Grade**
Absence of plaque	0
Up to 29% (usually one small plaque)	1
30 to 50% (1 medium plaque or several small plaques)	2
50 to 100% (1 large plaque or several plaques with at least 1 medium plaque)	3

Imputed plaque variables based on available information
The plaque index variable for some observations were flagged due to more number of plaques being noted on the CQI form compared to the IMT worksheet. Based on review of the CQI form, a minimal plaque index could be imputed and recorded in the following minimal imputed plaque index variables.

Variables	Label
pil_min	The minimum imputed plaque index of left side
pi_min	The minimum imputed plaque index
pir_min	The minimum imputed plaque index of right side

Equations for Created Carotid Variables:

- Mean IMT of average CCA (**MAVG_CCA**)
```
mavg_cca = (tcafr + tcanr + tcafl + tcanl) / 4;
```
- Mean IMT of average CCA – Right (**MAVGR_CCA)**
```
mavgr_cca = (tcafr + tcanr) / 2;
```
- Mean IMT of average CCA – Left (**MAVGL_CCA)**
```
mavgl_cca = (tcafl + tcanl) / 2;
```
- Max IMT of average CCA - Right (**MMAXR_CCA)**
```
mmaxr_cca = (tcxfr+tcxnr)/2;
```
- Max IMT of average CCA – Left (**MMAXL_CCA)**
```
mmaxr_cca = (tcxfr + tcxnr) / 2;
```
- Max IMT of average CCA **(MMAX_CCA**)
```
mmax_cca = (tcxfl + tcxnl + tcxfr + tcxnr) / 4;
```
- Minimum IMT of average CCA-Right (**MMINR_CCA**)
```
mminr_cca = (tcmfr + tcmnr) / 2;
```
- Minimum IMT of average CCA – Left (**MMINL_CCA)**
```
mminl_cca = (tcmfl + tcmnl) / 2;
```
- Minimum IMT of average CCA (**MMIN_CCA**)
```
mmin_cca = (tcmfl + tcmnl + tcmfr + tcmnr) / 4;
```
- Mean Average CCA Adventitial diameter **(ADAVG**)
```
ADAVG = (ADAVGr+ADAVGl) / 2;
```
- Mean minimum CCA Adventitial diameter (**ADMIN**)
```
ADMIN = (ADMINr+ADMINl) / 2;
```
- Mean maximum CCA Adventitial diameter (**ADMAX**)
```
ADMAX = (ADMAXr+ADMAXl) / 2;
```
- Lumen Diameter (**LUMEN**)
```
lumen = ((l4lenr1 + l4lenr2 + l4lenl1 + l4lenl2) / 4);
```
- Lumen Diameter -Right (**LUMENR**)
```
lumenr = ((l4lenr1 + l4lenr2) / 2);
```
- Lumen Diameter -Left (**LUMENL**)
```
lumenl = ((l4lenl1 +l4lenl2)/2);
```

Created plaque variables from the older version (10/13/2009**) of the URL Carotid IMT Worksheet**

- Right plaque index (**PIR**)
```
PIR = SUM (gra1r, gra2r, gra3r, grair, graer);
```
- Left plaque index (**PIL**)
```
PIL = SUM (gra1l, gra2l, gra3l, grail, grael);
```
- Plaque Index (**PI**)
```
PI = SUM(gra1r, gra2r, gra3r, grair, graer,
     gra1l, gra2l, gra3l, grail, grael);
```
- Presence or Absence of plaque (**PLAQUE**)
```
if PI > 0 then plaque = 1; else plaque = 0;
```

Created plaque variables from the current version (03/09/2010**) of the URL Carotid IMT Worksheet**

- Right plaque index (**PIR**)
```
if PVISUALR = 0 then PIR = 0;
   else PIR = SUM (gra1r, gra2r, gra3r, grair, graer);
```

- Left plaque index (**PIL**)

```
if PVISUALL = 0 then PIL=0;
   else PIL = SUM (gra1l, gra2l, gra3l, grail, grael);
```

- Plaque Index (**PI**)

```
PVIS = sum(PVISUALR, PVISUALL);
if PVIS = 0 then PI = 0;
if PVIS > 0 then PI = SUM(gra1r, gra2r, gra3r, grair, graer,
                          gra1l, gra2l, gra3l, grail, grael);
```

- Presence or Absence of plaque (**PLAQUE**)

```
if PI > 0 then Plaque = 1; else Plaque = 0
```