



FROM: Marston Youngblood, HCHS Coordinating Center

TO: HCHS/SOL Investigators

COPY: Jianwen Cai, Gerardo Heiss

DATE: March 4, 2019

RE: Distribution of HCHS/SOL Visit 2 Examination Investigator Use Database – Version 3 Release

123 W. FRANKLIN STREET
SUITE 450
CAMPUS BOX 8030
CHAPEL HILL, NC 27599-8030

T 919.962.6971
F 919.962.3265
www.csc.c.unc.edu

The HCHS/SOL Coordinating Center has prepared a full data release for all three years of the Visit 2 examinations. This is the full comprehensive release update of this data to the study investigators and the project office and NIH funding partners for Visit 2. The cut-off date for inclusion of information in this data release was through March 5, 2019 in order to process Echocardiography Reading Center quality control results to be integrated into the study database for all participant seen at the end of December 2017.

Summary of updates (details are at the end):

PART_DERV_V2 (Participant Derived Variables): 9 new variables and 17 updated/corrected variables.

ECHO (Echocardiography): Final data received by ECHO Reading Center.

PCE (Pregnancy Complications): 15 records dropped because were not valid pregnancy complications.

PCE_DERV: Changed conversion of pregnancy length from months to weeks. Updated 5 variables that use pregnancy length.

ALL FORMS. Identical to INV2 but we added variable FORM (language of administration).

The contents for this release include SAS datasets and following documentation:

1. HCHS/SOL Visit 2 Examination Database Overview Version 3.0
2. HCHS/SOL Visit 2 Examination Database Codebook Version 3.0
3. HCHS/SOL Visit 2 Derived Variable Dictionary Version 3.0
4. HCHS/SOL Visit 2 Pregnancy Complications Derived Variable Dictionary Version 3.0

As a reminder, investigator use data contain potentially identifying information and should be handled in a manner that ensures data security and confidentiality. These datasets must be stored on a secure server, so that authorized study team members access the data directly from either a local area network or remotely via a virtual private network in a secure fashion. At no time should investigator use datasets or subsets of the data be stored on laptop computers or portable data storage devices (e.g., memory sticks), nor should copies be distributed via email attachments. Making multiple copies for storage on stand-alone workstations is also prohibited, as the likelihood of data being hacked into is increased as a result. Each study PI is responsible for recording and tracking where the data are stored and who has access to the data at all times. Access to investigator use data should be granted to co-investigators or other study personnel at a PI's institution only for approved HCHS/SOL activities.

Please contact us at the Coordinating Center if you have any questions about the contents of this Visit 2 data distribution for exclusive use by the HCHS/SOL investigators and NIH.

Summary of Changes for Datasets and Data Values in the HCHS/SOL INV3 Data Release

1. Pregnancy Complication Datasets (PCE_V2_INV3, PCE_DERV_V2_INV3)

We re-calculated gestational age (PCE2A_WEEKS) and all derived variables.

- 10 months → 42 weeks
- 9 months → 40 weeks
- 8 months → 36 weeks
- 7 months → 32 weeks
- 6 months → 28 weeks

2. New Derived Variables

The participant derived variable dataset for Visit 2, Part_Derv_V2_INV3 contains additional variables that are documented in the dictionary that accompanies this release and they are profiled in the codebook.

YRSUS_V2	(Number of years in US visit 2)
HYPERTENSION2_AHA_V2	(Hypertension using new ACC/AHA Guidelines definition at visit 2)
HYPERTENSION_AHA_C5_V2	(5-level grouped hypertension (No HBP/Elevated/ treated HBP/ untreated Stage1 Hypertn /untreated stage2 Hypertn) at visit 2, New ACC/AHA Guideline)
STROKE_V2	(Prevalent Stroke, self report, Visit 2)
STROKE_TIA_V2	(Prevalent Stroke or TIA, self report, Visit 2)
CHD_SELF_V2	(Self-reported CHD Visit 2)
DIABETES_C4_V2	(4-level grouped diabetes (none/pre/treated/untreated) in V2)
DNA_AVAILABLE_V2	(DNA Available at Visit 2)
DIABETES4_V2	(Updated version of Diabetes2_V2)

3. Updated/Corrected Derived Variables

Derived variables that we updated or corrected in this release include the following:

YRSUS_C2_V2 - (2-level grouped years lived in the US (50 states) at Visit 2)

MUE 26a through 26p were set as 0 instead of missing for participants not taking any medications. (N=3645, out of 11623).

Impact on the existing derived variables from the medication survey recoding change in Ver. 3:

Updating/ corrected variables	Number of Changes from missing value to 0, out of N=11623 at Visit 2
DIABETES4_V2	20
HYPERTMED_SELF_V2	0
ELEVATED_BP_SELFMEDS_V2	1
IFG_NCEP_SELFMEDS_V2	29
IFG_IDF_SELFMEDS_V2	29
METS_NCEP2_V2	15
METS_IDF3_V2	0
DM_AWARE_V2	0
DM_TRT_V2**	3645 → of mue26 switched missing to 0
HIGH_TOTAL_CHOL2_V2	20
HYPERTENSION2_V2	1
HYPERTENSION_C4_V2	1
HYPERT_TREATMENT_V2	0
DYSLIPIDEMIA_C3_V2*	1282 from missing to 2 (dyslipidemia w/o treatment)
DYS_TCHDL_MED_V2	20
HYPERTENSION2_AHA_V2	1
HYPERTENSION_AHA_C5_V2	1

Notes:

*No positive category changes except Dyslipidemia_C3_v2 when recode for missing applied

**Dyslipidemia_C3_v2 has relatively large change in category 2 (Dyslipidemia W/O self-reported treatment). FYI, we do not have baseline version of Dyslipidemia_C3 released in visit 1.

**DM_TRT is the same as MUEA33c and DM_TRT_V2 is the same as MUE26c, and we normally use this variable within the category of Diabetics and no changes found in this case.