

Coding Manual of Curated Gen 3/NOS/Omni 2 dataset

Note:

1. This curated dataset contains information of Gen III, NOS, and Omni 2 participants
2. There are 4608 unique FHS participants within this dataset
3. Variables are generally grouped based on their origins (i.e., primary FHS dataset)
4. The format of the curated dataset is of long format, with distinct variable names that allow users to tell the difference for repeated measurements of data of similar nature.
5. For protocol descriptions and annotated forms of the primary datasets, refer to the respective documents of the individual source datasets.
6. Refer to Appendix for FHS core exam timeline (all cohorts)

Table of Contents

<u>Demographics</u>	3
<u>Anthropometric Measurements</u>	6
<u>Smoking</u>	7
<u>Blood Pressures, Heart Rate, Ventricular Hypertrophy and Antihypertensives</u>	8
<u>Blood glucose, Diabetes and Treatment</u>	9
<u>Lipid Panel and Hyperlipidemia Treatment</u>	11
<u>Creatinine</u>	12
<u>Neuropsychological assessments</u>	13
<u>Matching NP Dates with Core Exam Dates</u>	25
<u>Verified Outcomes</u>	
<u>Dementia and Related Variables</u>	26
<u>First Coronary Heart Disease (CHD) Event</u>	29
<u>First Stroke (CVA) Event</u>	30
<u>First Atrial Fibrillation (AF) Event</u>	31
<u>First Cancer and Related Variables</u>	32
<u>Survival Information</u>	35
<u>Center for Epidemiologic Studies Depression Scale (CES-D)</u>	36
<u>Neuropath Data and Availability</u>	42
<u>Appendix</u>	43
<u>Revisions</u>	44

Demographics

IDTYPE

Variable name: idtype

Description: Framingham Heart Study cohort identifier

IDTYPE and ID must be used together to form a unique study participant identification number or use variable “framid” instead

Code or Value	Description	Count
2	NOS	103
3	Gen III	4095
72	Omni 2	410

ID

Variable name: id

Description: Framingham Heart Study participant identifier within cohort

IDTYPE and ID must be used together to form a unique study participant identification number or use variable “framid” instead

Code or Value	Description
0 – 9999	Range of values

Note: To preserve confidentiality, the exact range for the ID is not reflected in the coding manual.

FRAMID

Variable name: Framid

Description: Framingham ID assigned to each participant

Code or Value	Description
20000 – 29999	NOS
30000 – 39999	Gen III
720000 – 729999	Omni 2

Note: To preserve confidentiality, the exact range for the ID is not reflected in the coding manual.

Demographics

SEX

Variable name: sex

Description: Gender

Code or Value	Description
1	Male
2	Female

RACE_CODE

Variable name: RACE_CODE

Description: RACE and ETHNICITY - Summary code string from combining primary codes given at different exams

Information extracted from "vr_raceall_2008_a_0712"

Code or Value	Description
W	White
B	Black
H	Hispanic or Latino
E	Not Hispanic or Latino
A	Asian
P	Native Hawaiian or Pacific Islander
N	American Indian or Alaskan Native
X	Asian Indian or Pacific Islander
R	Prefer not to answer
O	Other
.	Unknown

Note: RACE_CODE is constructed using primary codes W, B, H etc.

Race at different time points (exams) were combined to form RACE_CODE to indicate all races reported by participant over time.

Example1: RACE_CODE = 'EW' participant reported 'Not Hispanic or Latino' (H701) and 'White' (H702).

Example2: RACE_CODE = 'HH' participant reported 'Hispanic or Latino' (RACEATOMNI1) and 'Hispanic or Latino' (H700).

Example3: RACE_CODE = 'BEB' participant reported 'Black' (RACEATOMNI1), 'Not Hispanic or Latino' (H701) and 'Black' (H703).

Example4: RACE_CODE = 'ER' participant reported 'Not Hispanic or Latino' (H701) and 'Prefer not to answer' (H707).

Demographics

Character variable

DATE_CORE1 — DATE_CORE3

Variable name: date_core1 — date_core3

Description: Exam date of core exam 1 — 3 (MMDDYYYY)

Note: Date includes month, day and year

AGE_CORE1 — AGE_CORE3

Variable name: Age_core1 — Age_core3

Description: Age at core exam 1 — 3

EDU_CORE1 — EDU_CORE3

Variable name: Edu_core1 — Edu_core3

Description: Education at core exam 1 — 3 (If currently enrolled, the highest grade completed or degree received.)

Code or Value	Description
0	No schooling
1	Grades 1 – 8
2	Grades 9 – 11
3	Completed high school (12 th grade) or GED
4	Some college, but no degree
5	Technical school certificate
6	Associate degree (Junior college AA, AS)
7	Bachelor's degree (BA, AB, BS)
8	Graduate or professional degree (Master's, Doctorate, MD, etc.)
.	Prefer not to answer/Unknown

Note:

- For edu_core2 only, “Prefer not to answer” was coded as “9” and Unknown was coded as “.”.
- Education information is also collected during NP testing. Refer to variable “education_B1”, “education_B2” and “educg”

Anthropometric Measurements

HEIGHT_CORE1 — HEIGHT_CORE3

Variable name: Height_core1 — Height_core3

Description: Height at Core Exam 1 — 3 (inches, to next lower 1/4 inch)

WEIGHT_CORE1 — WEIGHT_CORE3

Variable name: Weight_core1 — Weight_core3

Description: Weight at Core Exam 1 — 3 (pounds)

BMI_CORE1 — BMI_CORE3

Variable name: BMI_core1 — BMI_core3

Description: Body mass index at Core Exam 1 — 3 (kg/m²)

HIP_CORE2 — HIP_CORE3

Variable name: Hip_core2 — Hip_core3

Description: Hip girth at Exam 2 — 3 (inches, to next lower 1/4 inch)

Note: Hip girth was not measured for Core Exam 1.

WAIST_CORE1 — WAIST_CORE3

Variable name: Waist_core1 — Waist_core3

Description: Waist girth at Core Exam 1 — 3 (inches, to next lower 1/4 inch)

Smoking

SMOKING_CORE1 — SMOKING_CORE3

Variable name: Smoking_core1 — Smoking_core3

Description: Current smoking status, Exam 1 — 3

Code or Value	Description
0	Not a current smoker
1	Current Smoker
.	Unknown

CPD_CORE1 — CPD_CORE3

Variable name: CPD_core1 — CPD_core3

Description: Number of cigarettes smoked per day, Exam 1 — 3

SBP_CORE1 — SBP_CORE3

Variable name: SBP_core1 — SBP_core3

Description: Average systolic blood pressure, Exam 1 — 3 (mmHg)

DBP_CORE1 — DBP_CORE3

Variable name: DBP_core1 — DBP_core3

Description: Average diastolic blood pressure, Exam 1 — 3 (mmHg)

VENT_RT_CORE1 — VENT_RT_CORE3

Variable name: Vent_rt_core1 — Vent_rt_core3

Description: Ventricular rate per minute by ECG, Exam 1 — 3

DLVH_CORE1 — DLVH_CORE3

Variable name: DLVH_core1 — DLVH_core3

Description: Definite left ventricular hypertrophy, Exam 1 — 3

Code or Value

Description

0	No
1	Yes
.	Unknown

HRX_CORE1 — HRX_CORE3

Variable name: HRX_core1 — HRX_core3

Description: Treated for hypertension, Exam 1 — 3

Code or Value

Description

0	No
1	Yes
.	Unknown

BG_CORE1 — BG_CORE3

Variable name: BG_core1 — BG_core3

Description: Blood glucose (includes fasting and non-fasting), Exam 1 — 3 (mg/dL)

FASTING_BG_CORE3 — FASTING_BG_CORE3

Variable name: Fasting_BG_core1 — Fasting_BG_core3

Description: Fasting (≥ 8 hours) blood glucose, Exam 1 — 3 (mg/dL)

DMRX_CORE1 — DMRX_CORE3

Variable name: DMRX_core1 — DMRX_core3

Description: Treated for diabetes, Exam 1 — 3

Code or Value	Description
0	No
1	Yes
.	Unknown

CURR_DM_CORE1 — CURR_DM_CORE3

Variable name: Curr_dm_core1 — Curr_dm_core3

Description: Current Diabetes status, Exam 1 — 3

Code or Value	Description
0	No
1	Yes
.	Unknown or did not attend Exam 1 — 3

HX_DM_CORE1 — HX_DM_CORE3

Variable name: Hx_dm_core1 — Hx_dm_core3

Description: History of diabetes, Exam 1 — 3

Code or Value	Description
0	No
1	Yes
.	Unknown or did not attend Exam 1 — 3

TC_CORE1 — TC_CORE3

Variable name: TC_core1 — TC_core3

Description: Total cholesterol, Exam 1 — 3 (mg/dL)

TRIGLYCERIDES_CORE1 — TRIGLYCERIDES_CORE3

Variable name: Triglycerides_core1 — Triglycerides_core3

Description: Triglycerides, Exam 1 — 3 (mg/dL)

CALC_LDL_CORE1 — CALC_LDL_CORE3

Variable name: Calc_LDL_core1 — Calc_LDL_core3

Description: Calculated LDL cholesterol, Exam 1 — 3 (mg/dL)

HDL_CORE1 — HDL_CORE3

Variable name: HDL_core1 — HDL_core3

Description: HDL-Cholesterol, Exam 1 — 3 (mg/dL)

LIPRX_CORE1 — LIPRX_CORE9

Variable name: LIPRX_core1 — LIPRX_core9

Description: Treated for lipids, Exam 1 — 3

Code or Value	Description
0	No
1	Yes
.	Unknown

Creatinine

CREATININE_CORE1 — CREATININE_CORE3

Variable name: Creatinine_core1 — Creatinine_core3

Description: Creatinine, Exam 1 — 3 (mg/dL)

EXAMDATE_NP1 — EXAMDATE_NP4

Variable name: examdate_NP1 — exam_NP4
Description: Date of the first — fourth NP exam (MMDDYYYY)

AGE_NP1 — AGE_NP4

Variable name: age_NP1 — age_NP4
Description: Age of participant at first — fourth NP exam

BATTERY_NP1 — BATTERY_NP4

Variable name: battery_NP1 — battery_NP4
Description: Form of the battery used for first — fourth NP exam

Code or Value	Description
1	Short form
2	Long form
.	Unknown

MARITAL_NP1 — MARITAL_NP4

Variable name: marital_NP1 — marital_NP4
Description: Current marital status for first — fourth NP exam

Code or Value	Description
1	Single
2	Married
3	Widowed
4	Divorced
5	Separated
6	Other
.	Unknown/Missing

HANDEDNESS_NP1 — HANDEDNESS_NP4

Variable name: handedness_NP1 — handedness_NP4
Description: Subject's handedness, NP Exam 1 — 4 (self-reporting)

Code or Value	Description
---------------	-------------

Neuropsychological assessment

1	Right
2	Left
3	Ambidextrous
.	Unknown/Missing

EMPLOYMENT_NP1 — EMPLOYMENT_NP4

Variable name: employment_NP1 — employment_NP4

Description: Employment status, NP exam 1 — 4

Code or Value	Description
0	Works full time
1	Works part-time
2	Not working due to disability
3	Retired, no longer works
4	Unemployed (expects to return to work)
5	Never worked for pay outside home
6	Volunteer
7	Other
10	Full-time student
11	Homemaker, expects to return to work
.	Unknown/Missing

EDUCATION_B1_NP1 — EDUCATION_B1_NP4

Variable name: education_b1_NP1 — education_b1_NP4

Description: Highest Education Level Completed (Collected from NP battery1), NP Exam 1 — 4

Code or Value	Description
0	None
1	4th grade or less
2	5th, 6th, or 7th grades
3	Grade school graduate; no high school
4	High school, did not graduate
5	High school diploma
6	College, did not graduate
7	College graduate
8	Postgraduate school
10	High school and business college, nursing school, music school, or art school only

.	Unknown
---	---------

EDUCATION_B2_NP1 — EDUCATION_B2_NP4

Variable name: education_b2_NP1 — education_b2_NP4

Description: Highest Education Level Completed (Collected from NP battery2), NP Exam 1 — 4

Code or Value	Description
0	None
1	4th grade or less
2	5th, 6th, or 7th grades
3	Grade school graduate; no high school
4	High school, did not graduate
5	High school diploma
6	GED
7	Some college course-work for credit (no certificate or degree)
8	Junior college – earned certificate or degree (e.g., AD)
9	Business college, nursing school, music school, art school
10	Technical School (post-HS) – earned certificate
11	Four Year College, partial
12	Four Year College, graduate
13	Graduate School, partial (no Masters degree)
14	Graduate School (Masters)
15	Graduate School (Doctorate)
.	Unknown

EDUCG_NP1 — EDUCG_NP4

Variable name: educg_NP1 — educg_NP4

Description: Education group (derived variable), NP Exam 1 — 4

Code or Value	Description
0	High school did not graduate
1	High school graduate
2	Some college

Neuropsychological assessment

3	College graduate
.	Unknown/Missing

OCCUP_B1_NP1 — OCCUP_B1_NP4

Variable name: occup_b1_NP1 — occup_b1_NP4

Description: Occupation from NP battery1, NP Exam 1 — 4

Code or Value	Description
1	Major professionals
2	Business managers
3	Administrative personnel
4	Clerical and sales workers
5	Skilled manual employees
6	Semi-skilled employees
7	Unskilled employed
8	Never employed

OCCUP_B2_NP1 — OCCUP_B2_NP4

Variable name: occup_b2_NP1 — occup_b2_NP4

Description: Occupation from NP battery2, NP Exam 1 — 4

Code or Value	Description
1	Professional, Technical and Related Occupations
2	Executive, Administrative, and Managerial Occupations
3	Sales Occupations
4	Administrative Support Occupations, Including Clerical
5	Precision Production, Craft, and Repair Occupations
6	Machine Operators, Assemblers, and Inspectors
7	Transportation and Material Moving Occupations
8	Handlers, Equipment Cleaners, Helpers, and Laborers
9	Service Occupations, Except Private Household

10	Never employed for pay (e.g., homemaker)
.	Unknown

LMI_NP1 — LMI_NP4

Variable name: lmi_NP1 — lmi_NP4

Description: Total number correct verbatim and paraphrases from “Logical Memories – Immediate Recall”, NP Exam 1 — 4 (Range: 0 – 23)

LMD_NP1 — LMD_NP4

Variable name: lmd_NP1 — lmd_NP4

Description: Total number correct verbatim and paraphrases from “Logical Memories – Delayed Recall”, NP Exam 1 — 4 (Range: 0 – 23)

LMR_NP1 — LMR_NP4

Variable name: lmr_NP1 — lmr_NP4

Description: Logical Memories – Recognition, NP Exam 1 — 4 (Range: 0 – 11)

LMI_JOE_NP1 — LMI_JOE_NP4

Variable name: lmi_joe_NP1 — lmi_joe_NP4

Description: Total number of correct verbatim and paraphrases from WMS-III “Logical Memory – Immediate Recall” (only for Joe Garcia), NP Exam 1 — 4 (Range: 0 – 23)

LMD_JOE_NP1 — LMD_JOE_NP4

Variable name: lmd_joe_NP1 — lmd_joe_NP4

Description: Total number of correct verbatim and paraphrases from WMS-III “Logical Memory – Delayed Recall” (only for Joe Garcia), NP Exam 1 — 4 (Range: 0 – 23)

LMR_JOE_NP1 — LMR_JOE_NP4

Variable name: lmr_joe_NP1 — lmr_joe_NP4

Description: Total number of correct verbatim and paraphrases from WMS-III “Logical Memory – Recognition” (only for Joe Garcia), NP Exam 1 — 4 (Range: 0 – 11)

LM_STORY_NP1 — LM_STORY_NP4

Variable name: LM_story_NP1 — LM_story_NP4

Description: Which logical memory story was administered? (NP Exam 1 — 4)

Code or Value	Description
1	Anna Thompson ONLY
2	Joe Garcia ONLY
3	Anna Thompson AND Joe Garcia
.	Neither story

VRI_NP1 — VRI_NP4

Variable name: vri_NP1 — vri_NP4

Description: Total points scored from “Visual Reproductions – Immediate Recall”, NP Exam 1 — 4 (Range: 0 – 14)

VRD_NP1 — VRD_NP4

Variable name: vrd_NP1 — vrd_NP4

Description: Total points scored from “Visual Reproductions – Delayed Recall”, NP Exam 1 — 4 (Range: 0 – 14)

VRR_NP1 — VRR_NP4

Variable name: vrr_NP1 — vrr_NP4

Description: Total points scored from “Visual Reproductions – Recognition”, NP Exam 1 — 4 (Range: 0 – 4)

PASI_NP1 — PASI_NP4

Variable name: pasi_NP1 — pasi_NP4

Description: Total Score from “Paired Associate Learning – Immediate Recall”, NP Exam 1 — 4 (Range 0 – 21)

PASI_E_NP1 — PASI_E1_NP4

Variable name: pasi_e_NP1 — pasi_e_NP4

Description: Easy Score from “Paired Associate Learning – Immediate Recall”, NP Exam 1 — 4 (Range: 0 – 18)

PASI_H_NP1 — PASI_H_NP4

Variable name: pasi_h_NP1 — pasi_h_NP4

Description: Hard Score from “Paired Associate Learning – Immediate Recall”, NP Exam 1 — 4 (Range: 0 – 12)

PASD_NP1 — PASD_NP4

Variable name: pasd_NP1 — pasd_NP4

Description: Total Score from “Paired Associate Learning – Delayed Recall”, NP Exam 1 — 4 (Range: 0 – 10)

PASD_E_NP1 — PASD_E_NP4

Variable name: pasd_e_NP1 — pasd_e_NP4

Description: Easy Score from “Paired Associate Learning – Delayed Recall”, NP Exam 1 — 4 (Range: 0 – 6)

PASD_H_NP1 — PASD_H_NP4

Variable name: pasd_h_NP1 — pasd_h_NP4

Description: Hard Score from “Paired Associate Learning – Delayed Recall”, NP Exam 1 — 4 (Range: 0 – 4)

PASR_NP1 — PASR_NP4

Variable name: pasr_NP1 — pasr_NP4

Description: Total Score from “Paired Associate Learning - Recognition”, NP Exam 1 — 4 (Range: 0 – 10)

DSF_NP1 — DSF_NP4

Variable name: dsf_NP1 — dsf_NP4

Description: Highest string from “Digits Forward Span”, NP Exam 1 — 4 (Range: 0 – 9)

DSB_NP1 — DSB_NP4

Variable name: dsb_NP1 — dsb_NP4

Description: Highest string from “Digits Backward Span”, NP Exam 1 — 4 (Range: 0 – 8)

TRAILSA_NP1 — TRAILSA_NP4

Variable name: trailsA_NP1 — trailsA_NP4

Description: Time in minutes to finish “Trail A”, NP Exam 1 — 4 – capped at 7 minutes (Range: 0.10 – 7.0)

TRAILSB_NP1 — TRAILSB_NP4

Variable name: trailsB_NP1 — trailsB_NP4
Description: Time in minutes to finish “Trail B”, NP Exam 1 — 4 — capped at 10 minutes (Range: 0.32 — 10.0)

SIM_NP1 — SIM_NP4

Variable name: sim_NP1 — sim_NP4
Description: Total raw score from “Similarities Test”, NP Exam 1 — 4 (Range: 0 — 26)

HVOT_NP1 — HVOT_NP4

Variable name: hvot_NP1 — hvot_NP4
Description: Total score from “Hooper Visual Organization Test”, NP Exam 1 — 4 (Range: 0 — 30)

BNT10_NP1 — BNT10_NP4

Variable name: bnt10_NP1 — bnt10_NP4
Description: Correct score without cue from 10 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 — 10)

BNT10_SEMANTIC_NP1 — BNT10_SEMANTIC_NP4

Variable name: bnt10_semantic_NP1 — bnt10_semantic_NP4
Description: Correct score with semantic cue from 10 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 — 10)

BNT10_PHONEMIC_NP1 — BNT10_PHONEMIC_NP4

Variable name: bnt10_phonemic_NP1 — bnt10_phonemic_NP4
Description: Correct score with phonemic cue from 10 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 — 10)

BNT30_NP1 — BNT30_NP4

Variable name: bnt30_NP1 — bnt30_NP4
Description: Correct score without cue from 30 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 — 30)

BNT30_SEMANTIC_NP1 — BNT30_SEMANTIC_NP4

Variable name: bnt30_semantic_NP1 — bnt30_semantic_NP4

Description: Correct score with semantic cue from 30 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 – 30)

BNT30_PHONEMIC_NP1 — BNT30_PHONEMIC_NP4

Variable name: bnt30_phonemic_NP1 — bnt30_phonemic_NP4

Description: Correct score with phonemic cue from 30 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 – 30)

BNT36_NP1 — BNT36_NP4

Variable name: bnt36_NP1 — bnt36_NP4

Description: Correct score without cue from 36 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 – 36)

BNT36_SEMANTIC_NP1 — BNT36_SEMANTIC_NP4

Variable name: bnt36_semantic_NP1 — bnt36_semantic_NP4

Description: Correct score with semantic cue from 36 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 – 36)

BNT36_PHONEMIC_NP1 — BNT36_PHONEMIC_NP4

Variable name: bnt36_phonemic_NP1 — bnt36_phonemic_NP4

Description: Correct score with phonemic cue from 36 items of “Boston Naming Test”, NP Exam 1 — 4 (Range: 0 – 36)

FINGTAPR_NP1 — FINGTAPR_NP4

Variable name: FingTapR_NP1 — FingTapR_NP4

Description: Mean of five trials of “Finger-tapping” of right hand, NP Exam 1 — 4

FINGTAPL_NP1 — FINGTAPL_NP4

Variable name: FingTapL_NP1 — FingTapL_NP4

Description: Mean of five trials of “Finger-tapping” of left hand, NP Exam 1 — 4

WRAT_NP1 — WRAT_NP4

Variable name: wrat_NP1 — wrat_NP4

Description: Raw Score from “Wide Range Achievement Test –Reading”, NP Exam 1 — 4 (Range: 15 – 57)

FAS_NP1 — FAS_NP4

Variable name: fas_NP1 — fas_NP4

Description: The total words (not including the animal part) from the Controlled Oral Word Association Test, NP Exam 1 — 4

FAS_ANIMAL_NP1 — FAS_ANIMAL_NP4

Variable name: fas_animal_NP1 — fas_animal_NP4

Description: The total words from the animal part of the Category Naming Test, NP Exam 1 — 4

Note: This variable only available on “NP battery2”

BD_NP1 — BD_NP4

Variable name: bd_NP1 — bd_NP4

Description: Total Score from “Block Design”, NP Exam 1 — 4 (Range: 0 – 26)

WAIS_NP1 — WAIS_NP4

Variable name: wais_NP1 — wais_NP4

Description: Total correct answers from “WAIS test”, NP Exam 1 — 4 (Range: 0 – 29)

CODING_CORRECT_NP1 — CODING_CORRECT_NP4

Variable name: coding_correct_NP1 — coding_correct_NP4

Description: Total correct from WAIS-IV Coding, NP Exam 1 — 4

CODING_INCORRECT_NP1 — CODING_INCORRECT_NP4

Variable name: coding_incorrect_NP1 — coding_incorrect_NP4

Description: Total incorrect from WAIS-IV Coding, NP Exam 1 — 4

INCIDENTAL_PAISED_NP1 — INCIDENTAL_PAISED_NP4

Variable name: Incidental_paired_NP1 — Incidental_paired_NP4

Description: Total correct from Paired Incidental Learning (Given the number from Coding, the participant had to draw the appropriate symbol associated with that number; each number is presented twice), NP Exam 1 — 4

INCIDENTAL_FREE_NP1 — INCIDENTAL_FREE_NP4

Variable name: Incidental_free_NP1 — Incidental_free_NP4

Description: Total correct from Free Recall Incidental Learning (The participant was asked to draw all the symbols they remember from the Coding task, without having to identify which number accompanied the symbol), NP Exam 1 — 4

MATH_CORRECT_NP1 — MATH_CORRECT_NP4

Variable name: Math_correct_NP1 — Math_correct_NP4
Description: Total correct responses for Math Fluency, NP Exam 1 — 4

MATH_INCORRECT_ADD_NP1 — MATH_INCORRECT_ADD_NP4

Variable name: Math_incorrect_add_NP1 — Math_incorrect_add_NP4
Description: Incorrect responses for addition problems, NP Exam 1 — 4

MATH_INCORRECT_SUB_NP1 — MATH_INCORRECT_SUB_NP4

Variable name: Math_incorrect_sub_NP1 — Math_incorrect_sub_NP4
Description: Incorrect responses for subtraction problems, NP Exam 1 — 4

MATH_INCORRECT_MUL_NP1 — MATH_INCORRECT_MUL_NP4

Variable name: Math_incorrect_mul_NP1 — Math_incorrect_mul_NP4
Description: Incorrect responses for multiplication problems, NP Exam 1 — 4

RIGHT_OPEN_NP1 — RIGHT_OPEN_NP4

Variable name: right_open_NP1 — right_open_NP4
Description: Seconds from balance test on right leg with eyes open (maximum time is set at 30 seconds), NP Exam 1 — 4

LEFT_OPEN_NP1 — LEFT_OPEN_NP4

Variable name: left_open_NP1 — left_open_NP4
Description: Seconds from balance test on left leg with eyes open (maximum time is set at 30 seconds), NP Exam 1 — 4

RIGHT_CLOSE_NP1 — RIGHT_CLOSE_NP4

Variable name: right_close_NP1 — right_close_NP4
Description: Seconds from balance test on right leg with eyes closed (maximum time is set at 30 seconds), NP Exam 1 — 4

LEFT_CLOSE_NP1 — LEFT_CLOSE_NP4

Variable name: left_close_NP1 — left_close_NP4

Description: Seconds from balance test on left leg with eyes closed (maximum time is set at 30 seconds), NP Exam 1 — 4

MRI352_NP1 — MRI352_NP4

Variable name: mri352_NP1 — mri352_NP4

Description: Tester's impression of participant's dementia status (Not verified), NP Exam 1 — 4

Code or Value	Description
0	Not demented
0.25	Not demented/MCI
0.5	Mild Cognitive Impairment
1	Mildly Demented
1.5	Mild – Moderately Demented
2	Moderately Demented
2.5	Moderately – Severely Demented
3	Severely Demented
5	Non-progressive impairment
.	Unknown

TEST_LANGUAGE_NP1 — TEST_LANGUAGE_NP4

Variable name: test_language_NP1 — test_language_NP4

Description: Language used during NP Exam 1 — 4

Code or Value	Description
1	English
2	Spanish
.	Unknown/Missing

Matching NP Dates with Core Exam Dates

NP_CLOSEST_EXAM_DATE1 — NP_CLOSEST_EXAM_DATE4

Variable name: np_closest_exam_date1 — np_closest_exam_date4
Description: FHS core exam date that was closest to the first — fourth NP test

NP_DAYS_APART1 — NP_DAYS_APART4

Variable name: np_days_apart1 — np_days_4
Description: The number of days between the FHS core exam date that was closest to the first — twelfth NP test and the first — fourth NP test date.

Note: A negative value means that the FHS core exam that was closest to the NP test occurred prior to the NP test, while a positive value means that the closest FHS core exam occurred after the NP test.

NP_EXAM_CYCLE1 — NP_EXAM_CYCLE4

Variable name: np_exam_cycle1 — np_exam_cycle4
Description: FHS core exam cycle that was closest to the first — fourth NP test

Code or Value	Description
1	Core Exam 1
2	Core Exam 2
3	Core Exam 3
.	Unknown

NORMAL_DATE

Variable name: normal_date

Description: Last date documented to be cognitively intact

Note: In event that the day and/or month of onset are uncertain, the day is default to 15 and the month is default to 6. The year of onset is always certain for all non-missing values.

IMPAIRMENT_DATE

Variable name: impairment_date

Description: Date of cognitive impairment onset

Note: In event that the day and/or month of onset are uncertain, the day is default to 15 and the month is default to 6. The year of onset is always certain for all non-missing values.

MILD_DATE

Variable name: mild_date

Description: Date of diagnosis of mild dementia

Note: In event that the day and/or month of onset are uncertain, the day is default to 15 and the month is default to 6. The year of onset is always certain for all non-missing values.

MODERATE_DATE

Variable name: moderate_date

Description: Date of diagnosis of moderate dementia

Note: In event that the day and/or month of onset are uncertain, the day is default to 15 and the month is default to 6. The year of onset is always certain for all non-missing values.

SEVERE_DATE

Variable name: severe_date

Description: Date of diagnosis of severe dementia

Note: In event that the day and/or month of onset are uncertain, the day is default to 15 and the month is default to 6. The year of onset is always certain for all non-missing values.

EDDD

Variable name: eddd

Description: Earliest documented date of Dementia

Note: This date variable is an estimated onset date for dementia using the earliest dates of various supporting resources and/or derived from earliest documented date variables of “mild_date”, “moderate_date” and “severe_dates”. In certain situations, the “EDDD” may precede those derived using “mild_date”, “moderate_date” and “severe_dates”.

DEMRV046

Variable name: demrv046

Description: Probable Dementia present

Code or Value	Description
0	No
1	Yes
.	Unknown

Note: This variable refers to all-cause dementia.

DEMRV103

Variable name: demrv103

Description: Dementia subtype

Code or Value	Description
0	None
1	Alzheimer’s disease without stroke
2	Alzheimer’s disease with stroke
3	Vascular dementia without Alzheimer’s disease
4	Mixed dementia type (Alzheimer’s disease + vascular dementia)
5	Frontotemporal dementia
6	Dementia with Lewy bodies
7	Dementia that does not fit any other Category (progressive)
8	Dementia that does not fit any other category (non-progressive)

Verified Outcomes – Dementia and Related Variables

9	Cognitive Impairment – No Dementia
10	Dementia – Uncertain
.	Unknown

Note: Value “9” refers to Cognitive impairment without Dementia. These subjects are verified to have mild cognitive impairment (MCI). However, it is also possible that participants, who had not yet been screened by the dementia consensus panel, may also suffer from mild cognitive impairment. Hence, if the research outcome of interest is mild cognitive impairment, it is advisable to also consider the psychometric results of the study population as part of the outcome evaluation.

FIRST_CHD_DATE

Variable name: first_chd_date

Description: Date of first Coronary Heart Disease (CHD) Event

Note: This variable is derived from the Sequence of Event (SoE) dataset, using the following values in (1, 2, 3, 4, 5, 6, 7, 8, 21, 22, 23, 24) for the variable “EVENT”

Event Values in SoE Dataset Description

1	MI recognized, with diagnostic ECG
2	MI recognized, without diagnostic ECG, with enzymes and history
3	MI recognized, without diagnostic ECG, with autopsy evidence, new event (see also code 9)
4	MI unrecognized, silent
5	MI unrecognized, not silent
6	AP, first episode only
7	CI, definite by both history and ECG
8	Questionable MI at exam 1
21	Death, CHD sudden, within 1 hour
22	Death, CHD, 1 - 23 hours, nonsudden
23	Death, CHD, 24-47 hours, non sudden
24	Death, CHD, 48 hours or more, non sudden
.	No record in SoE dataset

Abbreviations:

MI – Myocardial Infarction

ECG – Electrocardiogram

AP – Angina Pectoris

CI – Coronary Insufficiency

FIRST_STROKE_DATE

Variable name: first_stroke_date

Description: Date of first Stroke (CVA) Event

Note: This variable is derived from the Sequence of Event (SoE) dataset, using the following values in (10, 11, 13, 14, 15, 16, 17, 19) for the variable “EVENT”

Event Values in SoE Dataset Description

10	Definite CVA at exam 1, but questionable type
11	ABI (Atherothrombotic Infarction of brain)
13	Cerebral embolism
14	Intracerebral hemorrhage
15	Subarachnoid hemorrhage
16	Other CVA
17	CVA, definite CVA, type unknown
19	Questionable CVA at exam 1
.	No record in SoE dataset

Abbreviations:

CVA – Cerebrovascular Accident

ABI – Atherothrombotic Infarction of Brain

Verified Outcomes – First Atrial Fibrillation (AF) Event

FIRST_AF_DATE

Variable name: first_af_date

Description: Date for first Atrial Fibrillation (AF) Event

Note: This variable is derived from the verified Atrial Fibrillation (AF) dataset, using the values in (1, 2, 3) for the variable “RHYTHM”

Rhythm Values in AF Dataset	Description
-----------------------------	-------------

1	Atrial fibrillation
2	Atrial flutter
3	Atrial fibrillation / atrial flutter
.	No record in AF dataset

FIRST_CANCER_DATE

Variable name: first_cancer_date

Description: First diagnosis date of tumor or cancer

Note: This variable is derived from the verified Cancer dataset, using the variable “D_DATE”

TOPO_CANCER

Variable name: topo_cancer

Description: Topography - describes site - corresponds to first three digits of topography code WHO ICD-O

Note: This variable is derived from the verified Cancer dataset, using the variable “TOPO”

Rhythm Values in AF Dataset Description

140	Lip
141	Tongue
142	Major salivary gland
143	Gum
144	Floor of mouth
145	Other and unspecified parts of mouth
146	Oropharynx
147	Nasopharynx
148	Hypopharynx
149	Pharynx and ill-defined sites in lip, oral cavity and pharynx
150	Esophagus
151	Stomach
152	Small intestine
153	Colon
154	Rectum, rectosigmoid junction, anal canal and anus, not otherwise specified
155	Liver and intrahepatic bile ducts
156	Gallbladder and extrahepatic bile ducts
157	Pancreas
158	Retroperitoneum and peritoneum
159	Other and ill-defined sites within digestive organs and peritoneum
160	Nasal cavities, accessory sinuses, middle ear and inner ear
161	Larynx
162	Trachea bronchus, and lung
163	Pleura

Verified Outcomes – First Cancer and Related Variables

164	Thymus, heart and mediastinum
165	Other and ill-defined sites within respiratory system and intrathoracic organs
169	Hematopoietic and reticuloendothelial systems
170	Bones, joints and articular cartilage
171	Connective, subcutaneous and other soft tissues
173	Skin (excludes skin of labia majora, skin of vulva, skin of penis and skin of scrotum)
174	Female breast (excludes skin of breast)
175	Male breast (excludes skin of breast)
179	Uterus, not otherwise specified
180	Cervix uteri
181	Placenta
182	Corpus uteri or endometrium
183	Ovary, fallopian tube and broad ligament
184	Other and unspecified female genital organs
185	Prostate gland
186	Testis
187	Penis and other male genital organs
188	Urinary bladder
189	Kidney and other urinary organs
190	Eye and lacrimal gland
191	Brain
192	Other and unspecified parts of the nervous system (excludes 4 peripheral nerves, sympathetic and parasympathetic nerves and ganglia)
193	Thyroid gland
194	Other endocrine glands
195	Other ill-defined sites
196	Lymph nodes
199	Unknown primary site
.	No record in Cancer dataset

LOCATION_CANCER

Variable name: location_cancer

Description: Location - describes subdivision of site - corresponds to fourth digit of topography code WHO ICD-O

Note: This variable is derived from the verified Cancer dataset, using the variable “LOCATION”

Code or Value	Description
---------------	-------------

Verified Outcomes – First Cancer and Related Variables

0 – 9	Range of values
-------	-----------------

HIST_CANCER

Variable name: hist_cancer

Description: Histology or morphology - describes cell histopathology - corresponds to four digit morphology code WHO ICD-O

Note: This variable is derived from the verified Cancer dataset, using the variable “HIST”

Code or Value	Description
8000 – 9990	Range of values

Survival Information

LAST_CONTACT_DATE

Variable name: Last_contact_date

Description: Date of last contact (This is the last time the participant was known to be alive)

Note: This date is censored to be the last day of 2019 if the last contact date was after the last day of 2019.

DEATH_DATE

Variable name: Death_date

Description: Date of death

CESD_CORE1_1 — CESD_CORE3_1

Variable name: CESD_core1_1 — CESD_core3_1

Description: CESD: I was bothered by things that usually don't bother me (at exam 1 — 3)

Code or Value	Description
0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_2 — CESD_CORE3_2

Variable name: CESD_core1_2 — CESD_core3_2

Description: CESD: I did not feel like eating: my appetite was poor (at exam 1 — 3)

Code or Value	Description
0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_3 — CESD_CORE3_3

Variable name: CESD_core1_3 — CESD_core3_3

Description: CESD: I felt that I could not shake off the blues, even with the help of family and friends (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_4 — CESD_CORE3_4

Variable name: CESD_core1_4 — CESD_core3_4

Description: CESD: I felt that I was just as good as other people (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_5 — CESD_CORE3_5

Variable name: CESD_core1_5 — CESD_core3_4

Description: CESD: I had trouble keeping my mind on what I was doing (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_6 — CESD_CORE3_6

Variable name: CESD_core1_6 — CESD_core3_6

Description: CESD: I felt depressed (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_7 — CESD_CORE3_7

Variable name: CESD_core1_7 — CESD_core3_7

Description: CESD: I felt that everything I did was an effort (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_8 — CESD_CORE3_8

Variable name: CESD_core1_8 — CESD_core3_8

Description: CESD: I felt hopeful about the future (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_9 — CESD_CORE3_9

Variable name: CESD_core1_9 — CESD_core3_9

Description: CESD: I thought my life had been a failure (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_10 — CESD_CORE3_10

Variable name: CESD_core1_10 — CESD_core3_10

Description: CESD: I felt fearful (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_11 — CESD_CORE3_11

Variable name: CESD_core1_11 — CESD_core3_11

Description: CESD: My sleep was restless (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days

Center for Epidemiologic Studies Depression Scale (CES-D)

.	Unknown
---	---------

CESD_CORE1_12 — CESD_CORE3_12

Variable name: CESD_core1_12 — CESD_core3_12

Description: CESD: I was happy (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_13 — CESD_CORE3_13

Variable name: CESD_core1_13 — CESD_core3_13

Description: CESD: I talked less than usual (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_14 — CESD_CORE3_14

Variable name: CESD_core1_14 — CESD_core3_14

Description: CESD: I felt lonely (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_15 — CESD_CORE3_15

Variable name: CESD_core1_15 — CESD_core3_15

Description: CESD: People were unfriendly (at exam 1 — 3)

0	<1 day
---	--------

Center for Epidemiologic Studies Depression Scale (CES-D)

1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_16 — CESD_CORE3_16

Variable name: CESD_core1_16 — CESD_core3_16

Description: CESD: I enjoyed life (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_17 — CESD_CORE3_17

Variable name: CESD_core1_17 — CESD_core3_17

Description: CESD: I had crying spells (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_18 — CESD_CORE3_18

Variable name: CESD_core1_18 — CESD_core3_18

Description: CESD: I felt sad (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_19 — CESD_CORE3_19

Variable name: CESD_core1_19 — CESD_core3_19

Center for Epidemiologic Studies Depression Scale (CES-D)

Description: CESD: I felt that people disliked me (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_CORE1_20 — CESD_CORE3_20

Variable name: CESD_core1_20 — CESD_core3_20

Description: CESD: I could not "get going" (at exam 1 — 3)

0	<1 day
1	1 – 2 days
2	3 – 4 days
3	5 – 7 days
.	Unknown

CESD_TOTAL_CORE1 — CESD_TOTAL_CORE3

Variable name: CESD_total_core1 — CESD_total_core3

Description: CESD total score at exam 1 — 3

Note: The total CES-D scores are calculated for participants who answered all 20 questions.

NPBRAAK

Variable name: NPBRAAK

Description: Braak stage for neurofibrillary degeneration (B score). Use standard blocks to assign phase.

Code or Value	Description
0	Stage 0: AD-type neurofibrillary degeneration not present (B0)
1	Stage I (B1)
2	Stage II (B2)
3	Stage III (B3)
4	Stage IV (B4)
5	Stage V (B5)
6	Stage VI (B6)

NPNIT

Variable name: NPNIT

Description: Derived NIA-Reagan Criteria

Code or Value	Description
1	High likelihood dementia due to AD
2	Intermediate likelihood dementia due to AD
3	Low likelihood dementia due to AD
4	Criteria not met

NEUROPATH_AVAIL

Variable name: neuropath_avail

Description: Availability of brain autopsy data

Code or Value	Description
0	No
1	Yes

Appendix

Exam Dates and Age Ranges as of 09/14/2016				
Original Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	1948 - 1953	28 - 74	44	5209
Exam 2	1950 - 1955	31 - 65	46	4792
Exam 3	1952 - 1956	32 - 67	48	4416
Exam 4	1954 - 1958	34 - 69	50	4541
Exam 5	1956 - 1960	37 - 70	52	4421
Exam 6	1958 - 1963	38 - 72	54	4259
Exam 7	1960 - 1964	40 - 74	55	4191
Exam 8	1962 - 1966	42 - 76	57	4030
Exam 9	1964 - 1968	44 - 78	59	3833
Exam 10	1966 - 1970	46 - 80	61	3595
Exam 11	1968 - 1971	49 - 81	62	2955
Exam 12	1971 - 1974	50 - 83	64	3261
Exam 13	1972 - 1976	53 - 85	66	3133
Exam 14	1975 - 1978	55 - 88	68	2871
Exam 15	1977 - 1979	57 - 89	69	2632
Exam 16	1979 - 1982	59 - 91	70	2351
Exam 17	1981 - 1984	61 - 93	72	2179
Exam 18	1983 - 1985	63 - 94	74	1825
Exam 19	1985 - 1988	65 - 96	75	1541
Exam 20	1986 - 1990	67 - 97	77	1401
Exam 21	1988 - 1992	69 - 99	79	1319
Exam 22	1990 - 1994	72 - 101	80	1166
Exam 23	1992 - 1996	73 - 101	81	1026
Exam 24	1995 - 1998	76 - 103	83	831
Exam 25	1997 - 1999	78 - 104	84	703
Exam 26	1999 - 2001	79 - 103	86	558
Exam 27	2002 - 2003	82 - 104	87	414
Exam 28	2004 - 2005	84 - 104	89	303
Exam 29	2006 - 2007	85 - 102	91	218
Exam 30	2008 - 2010	88 - 102	92	141
Exam 31	2010 - 2011	90 - 99	92	91
Exam 32	2012 - 2014	93 - 106	96	40
Offspring Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	1971 - 1975	5 - 70	36	5124
Exam 2	1979 - 1983	17 - 77	44	3863
Exam 3	1983 - 1987	18 - 77	48	3873
Exam 4	1987 - 1991	22 - 81	52	4019
Exam 5	1991 - 1995	26 - 84	55	3799
Exam 6	1995 - 1998	29 - 86	59	3532
Exam 7	1998 - 2001	33 - 90	62	3539
Exam 8	2005 - 2008	40 - 93	67	3021
Exam 9	2011 - 2014	46 - 98	71	2430
Generation 3 Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	2002 - 2005	19 - 72	40	4095
Exam 2	2008 - 2011	24 - 78	47	3411
Exam 3	2016 - 2019			
New Offspring Spouse Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	2003 - 2005	47 - 85	65	103
Exam 2	2008 - 2011	53 - 90	70	68
Exam 3	2016 - 2019			
Omni 1 Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	1994 - 1998	27 - 78	52	506
Exam 2	1999 - 2001	33 - 83	55	405
Exam 3	2007 - 2008	39 - 89	62	298
Exam 4	2011 - 2014	44 - 88	66	301
Omni 2 Cohort				
Exam	Exam Date Range	Age Range	Mean Age	Attendees
Exam 1	2003 - 2005	20 - 80	43	410
Exam 2	2009 - 2011	25 - 85	48	321
Exam 3	2016 - 2019			

Revisions	
08/03/2018	Original Draft <ul style="list-style-type: none"> • Coding Manual of Master_Dataset_2372_080318 • Dataset: Master_Dataset_2372_080318
08/13/2018	Revision of 08/03/2018 draft <ul style="list-style-type: none"> • Minor typo/grammatical/format corrections • Addition of “Derived Variables” section with SAS code embedded • Added information to cover page of manual • Amended date of “Last Updated” • Added user advice to “Neuropsychological Test”, “Digital Clock Drawing Test” and “MRI Brain Scans” sections • Added user advice to the variable Last_Contacted_Date under “Verified Survival Status (Alive/Dead)” • Coding manual has been renamed as “Coding Manual of Master_Dataset_2372_081318” • Dataset has been renamed as “Master_Dataset_2372_081318”
08/21/2018	Revision of 08/13/2018 draft <ul style="list-style-type: none"> • Minor typo/grammatical/format corrections • Coding manual has been renamed as “Coding Manual of Master_Dataset_2372_082118” • Dataset has been renamed as “Master_Dataset_2372_082118”
08/30/2018	Revision of 08/21/2018 draft <ul style="list-style-type: none"> • Added information regarding interpretation of the range values of the “Neuropsychological Test” section in the introductory summary • Minor typo/grammatical/format corrections • Coding manual has been renamed as “Coding Manual of Master_Dataset_2372_082118 (Last Updated 083018)” • No change to dataset or dataset name
Feb 2021	Revision of 08/30/2018 draft <ul style="list-style-type: none"> • Added variables “idtype” and “id”

	<ul style="list-style-type: none"> • Dataset has been renamed as "Curated_cohort2372_022021" • Available in both SAS and csv files
Nov 2021	<p>Revision of Feb 2021</p> <ul style="list-style-type: none"> • Added selected variables from FHS core exam 3 • Added dementia review data • Updated neuropsychological summary scores and date-related variables up to year 2019 • Dataset has been renamed as "curated2372_11_2021" • Available in both SAS and csv files
Feb 2023	<p>Revision of Nov 2021 draft</p> <ul style="list-style-type: none"> • Included data up to year 2022 • Simplify coding manual format to reduce the length • Renamed all NP test related variables using suffix "_NP" for clearer annotation • Added more clinical outcomes and related variables • Added neuropath section • Removed selected medications and biomarker data till harmonized variables are created • Typo/grammatical/format corrections • Provided SAS code to include list of requested datasets to generate similar curated dataset format