



Atherosclerosis Risk in Communities Study

Cohort Exam Visit 6 NCS

V2_V6_CNF Derived Variable Dictionary (v.1)

August 2019

Prepared by the Collaborative Studies Coordinating Center

ARIC V2_V6_CNF Derived Variable Dictionary

Table of Contents

New or Changed from PREVIOUS Distribution	3
1. Overview	4
2. Administrative.....	4
2.1 SUBJECTID (ARIC Subject ID (CIR))	4
2.2 ID (ARIC ID - same as SUBJECTID).....	4
2.3 VTYPE (Visit type).....	5
2.4 VDATE (Visit date)	5
3. Neurocognitive Z Scores.....	6
3.1 ZSCORE_DWR (Z score of delayed word recall)	6
3.2 ZSCORE_DSS (Z score of digit symbol substitution)	6
3.3 ZSCORE_WFT (Z score of word fluency test)	7
3.4 MISSINGTESTS (Number of missing neurocognitive tests)	8
3.5 GLOBZ_NCTS (Mean of DWR, DSS, and WFT z scores).....	8
4. Factor Scores from the CSCC	8
4.1 GLOBALFS1 (Global cognition factor score from categorized data (ver1)).....	9
4.2 LANGUAGEFS1 (Language domain factor score from categorized data (ver1))	10
4.3 EXECFUNCFS1 (Executive functioning domain factor score from categorized data (ver1))	10
4.4 MEMORYFS1 (Memory domain factor score from categorized data (ver1))	11
5. References	11

NEW OR CHANGED FROM PREVIOUS DISTRIBUTION

This table describes the changes to the last published V2_V6_CNF dictionary. As the dataset undergoes modifications, this table will describe the updates made to the previously distributed dataset.

Modification Date	Variable Name	Reason(s) for Change

1. OVERVIEW

The V2_V6_CNF_yymmdd dataset contains 41,159 records from 14,529 participants who completed one or more neurocognitive assessments between Visit 2 (1990-1992) and Visit 6 (2016-2017). The dataset utilizes a long format in which each participant has multiple records. Each record represents a specific visit or sub-study during which the participant completed an assessment. The purpose of this dataset is to provide ARIC collaborators with a set of variables that can be used to examine risk and protective factors associated with cognitive change over time.

The dataset naming conventions are as follows: The dataset name retains the retrieval date (ex: V2_V6_CNF_190719) until the dataset is considered final and frozen. After a dataset is frozen, the retrieval date is dropped from the dataset name (ex: V2_V6_CNF). The first two characters refer to the earliest visit included in the dataset. The next two characters indicate the last visit in the dataset. The variable naming convention is similar. Across-visit variables have identical names. The last digit in the variable name identifies the definition version of a variable.

Details about the administration of neurocognitive measures can be found in ARIC Manual 17. The construction of factor scores based on these neurocognitive measures is described in ARIC Manual 30.

2. ADMINISTRATIVE

Administrative variables are used to identify specific participants as well as the visit and date on which the participant completed a neurocognitive assessment.

2.1 SUBJECTID (ARIC Subject ID (CIR))

Type: Character; length: \$7.

2.2 ID (ARIC ID - same as SUBJECTID)

Description: The historical participant identifier from visits 1-4 is ID. The value of ID is the same value as SUBJECTID. Use ID when merging visit 6/NCS stage 1 data with datasets from previous visits necessary for longitudinal analyses.

Type: Character; length: \$7.

Algorithm: ID=SUBJECTID.

Source variable(s): SUBJECTID

2.3 VTYPE (Visit type)

Description: A variable created to indicate the specific visit at which an assessment was administered. Neurocognitive tests were administered at Visit 2 (1990-1992), Visit 3 (1993-1995), Visit 4 (1996-1998), Visit 5 (2011-2013), and Visit 6 (2016-2017). Neurocognitive tests were also administered during two MRI sub-studies ('BRAIN' from 2004 to 2006 and 'CarMRI' from 2005 to 2006).

Type: Character; length: \$6.

Algorithm: If ID is in [CNFA] then VTYPE='V2'.
If ID is in [CNFB04] then VTYPE='V3'.
If ID is in [CNFC04] then VTYPE='V4'.
If ID is in [NCS Visit 5] then VTYPE='V5NCS'.
If ID is in [NCS Visit 6] then VTYPE='V6NCS'.
If ID is in [CNF_IA1] then VTYPE='BRAIN'.
If ID is in [CNFF] then VTYPE='CarMRI'.

Source variable(s): ID

2.4 VDATE (Visit date)

Description: A variable created to indicate the specific date on which a neurocognitive test was administered.

Type: Date

Algorithm: If VTYPE='V2' then VDATE=[DERIVE2_10]V2DATE21.
If VTYPE='V3' then VDATE=[DERIVE37]V3DATE31.
If VTYPE='V4' then VDATE=[DERIVE46]V4DATE41.
If VTYPE='V5NCS' then VDATE=[NCS Visit 5]NCS0A.
Else if VTYPE='V5NCS' and VDATE is missing then VDATE=[DERIVE52]V5DATE51.
If VTYPE='V6NCS' then VDATE=[NCS Visit 6]NCS0A.
Else if VTYPE='V6NCS' and VDATE is missing then VDATE=[DERIVE61]V6DATE61.
If VTYPE='Brain' then VDATE=[CNF_IA1]CNFD97.
If VTYPE='CarMRI' then VDATE=[CNFF]CNFF5.

Source variable(s): VTYPE, V2DATE21, V3DATE31, V4DATE41, NCS0A, V5DATE51, V6DATE61, CNFD97, CNFF5

3. NEUROCOGNITIVE Z SCORES

During Visit 2, three neurocognitive tests were administered to each participant. These tests included the Delayed Word Recall, the Digit Symbol Substitution, and the Word Fluency Test which respectively measured memory, executive functioning, and language ability. The tests were re-administered at Visits 3 through 6 as well as during two MRI sub-studies (VTYPE='BRAIN' and VTYPE='CarMRI'). The raw scores of these tests were standardized to Visit 2 by calculating the mean ($mean_{v2}$) and standard deviation (sd_{v2}) at Visit 2, subtracting $mean_{v2}$ from all values, and dividing by sd_{v2} .

3.1 ZSCORE_DWR (Z score of delayed word recall)

Description: A z score version of the Delayed Word Recall standardized to Visit 2. The score ranges from -4.33 to 2.22.

Type: Numeric

Algorithm: If VTYPE='V2' then DWRSCORE=[CNFA]CNFA01.
If VTYPE='V3' then DWRSCORE=[CNFB04]CNFB1.
If VTYPE='V4' then DWRSCORE=[CNFC04]CNFC1.
If VTYPE='V5NCS' then DWRSCORE=[NCS Visit 5]NCS3B.
If VTYPE='V6NCS' then DWRSCORE=[NCS Visit 6]NCS3B.
If VTYPE='Brain' then DWRSCORE=[CNF_IA1]CNFD5.
If VTYPE='CarMRI' then DWRSCORE=[CNFF]CNFF2.
If DWRSCORE <0 or >10 then DWRSCORE is set to missing.

If VTYPE='V2' then calculate the mean of DWRSCORE as $DWRSCORE_{mean_{v2}}$ and the standard deviation of DWRSCORE as $DWRSCORE_{sd_{v2}}$. Calculate ZSCORE_DWR for all visits as $ZSCORE_DWR = (DWRSCORE - DWRSCORE_{mean_{v2}}) / DWRSCORE_{sd_{v2}}$.

Source variable(s): VTYPE, CNFA01, CNFB1, CNFC1, NCS3B, CNFD5, CNFF2

3.2 ZSCORE_DSS (Z score of digit symbol substitution)

Description: A z score version of the Digit Symbol Substitution standardized to Visit 2. The score ranges from -3.13 to 3.40.

Type: Numeric

Algorithm: If VTYPE='V2' then DSSSCORE=[CNFA]CNFA02.
If VTYPE='V3' then DSSSCORE=[CNFB04]CNFB2.
If VTYPE='V4' then DSSSCORE=[CNFC04]CNFC2.
If VTYPE='V5NCS' then DSSSCORE=[NCS Visit 5]NCS2B.

If VTYPE='V6NCS' then DSSSCORE=[NCS Visit 6]NCS2B.
If VTYPE='Brain' then DSSSCORE=[CNF_IA1]CNFD3.
If VTYPE='CarMRI' then DSSSCORE=[CNFF]CNFF3.

If VTYPE='V2' then calculate the mean of DSSSCORE as DSSSCOREmean_{V2} and the standard deviation of DSSSCORE as DSSSCOREsd_{V2}. Calculate ZSCORE_DSS for all visits as $ZSCORE_DSS = (DSSSCORE - DSSSCOREmean_{V2}) / DSSSCOREsd_{V2}$.

Source variable(s): VTYPE, CNFA02, CNFB2, CNFC2, NCS2B, CNFD3, CNFF3

3.3 ZSCORE_WFT (Z score of word fluency test)

Description: A z score version of the Word Fluency Test standardized to Visit 2. The score ranges from -2.66 to 3.37.

Type: Numeric

Algorithm:
If VTYPE='V2' then WFTSCORE=[CNFA]CNFA04.
If VTYPE='V3' then WFTSCORE=[CNFB04]CNFB4.
If VTYPE='V4' then WFTSCORE=[CNFC04]CNFC4.
If VTYPE='V5NCS' then WFTSCORE=[NCS Visit 5]NCS5E.
If VTYPE='V6NCS' then WFTSCORE=[NCS Visit 6]NCS5E.
If VTYPE='Brain' then WFTSCORE=[CNF_IA1]SUM (CNFD7A, CNFD7B, CNFD7C, CNFD9A, CNFD9B, CNFD9C, CNFD11A, CNFD11B, CNFD11C) if none are missing.
If VTYPE='CarMRI' then WFTSCORE=[CNFF]CNFF4.
If WFTSCORE <0 or >75 then WFTSCORE is set to missing.

If VTYPE='V2' then calculate the mean of WFTSCORE as WFTSCOREmean_{V2} and the standard deviation of WFTSCORE as WFTSCOREsd_{V2}. Calculate ZSCORE_WFT for all visits as $ZSCORE_WFT = (WFTSCORE - WFTSCOREmean_{V2}) / WFTSCOREsd_{V2}$.

Source variable(s): VTYPE, CNFA04, CNFB4, CNFC4, NCS5E, CNFD7A, CNFD7B, CNFD7C, CNFD9A, CNFD9B, CNFD9C, CNFD11A, CNFD11B, CNFD11C, CNFF4

3.4 MISSINGTESTS (Number of missing neurocognitive tests)

Description: A variable created to indicate whether the Delayed Word Recall, the Digit Symbol Substitution, or the Word Fluency Test was missing from a specific visit or sub-study. The count ranges from 0 to 3.

Type: Numeric

Algorithm: Count of missing from ZSCORE_DWR, ZSCORE_DSS, and ZSCORE_WFT.

Source variable(s): ZSCORE_DWR, ZSCORE_DSS, and ZSCORE_WFT

3.5 GLOBZ_NCTS (Mean of DWR, DSS, and WFT z scores)

Description: A score of global cognition calculated by computing the mean from the z score versions of the Delayed Word Recall, the Digit Symbol Substitution, and the Word Fluency Test administered during a specific visit or sub-study. The score ranges from -3.37 to 2.58.

Type: Numeric

Algorithm: If MISSINGTESTS=0 then GLOBZ_NCTS=mean(ZSCORE_DWR, ZSCORE_DSS, ZSCORE_WFT).

Source variable(s): MISSINGTESTS, ZSCORE_DWR, ZSCORE_DSS, ZSCORE_WFT

4. FACTOR SCORES FROM THE CSCC

At Visit 2 and Visit 4 three neurocognitive tests were administered. These included the Delayed Word Recall, the Digit Symbol Substitution, and the Word Fluency Test. During one of the MRI sub-studies (VTYPE='BRAIN'), eight neurocognitive tests were administered. The five additional tests were the Logical Memory Test (Wechsler, 1987), Incidental Learning (Ryan & Lopez, 2001), Animal Naming Score (Benton & Hamsher, 1976), Trail Making Test A (Reitan, 1958), and Trail Making Test B (Reitan, 1958). At Visit 5 and Visit 6, the Boston Naming Test (Williams, Mack, & Henderson, 1989) and the Digit Span Backwards (Wechsler, 1987) were added to the test battery. Raw scores from each neurocognitive test were discretized into ten or fewer categories. A categorical confirmatory factor analysis (Gross, et al., 2015) was utilized to compute a global cognition factor score for each participant. Separate factor scores for the cognitive domains of language, executive functioning, and memory (Hayden, et al., 2011; Park, et al., 2012; Rawlings, et al., 2016; Siedlecki, et al., 2010) were estimated. The ARIC cohort was treated as a single group in the factor analysis. Differential item functioning across subgroups was **not** used to compute the factor scores. Each factor

score was standardized to Visit 5 by calculating the mean (mean_{V5}) and standard deviation (sd_{V5}) at Visit 5, subtracting mean_{V5} from all values, and dividing by sd_{V5} . Additional details about the creation of factor scores can be found in ARIC Manual 30.

4.1 GLOBALFS1 (Global cognition factor score from categorized data (ver1))

Description: A factor score of global cognition estimated from a categorical confirmatory factor analysis of the Delayed Word Recall, Digit Symbol Substitution, Word Fluency Test, Logical Memory Test, Incidental Learning, Animal Naming Score, Trail Making Test A, Trail Making Test B, Boston Naming Test, and Digit Span Backwards. A factor score was computed for Visits 2, 4, and 5 as well as one of the MRI sub-studies (VTYPE='BRAIN') based on discretized versions of each neurocognitive test. The factor loadings and thresholds from Visit 5 were integrated as fixed parameters into a model that generated factor scores for Visit 6. Factor scores range from -2.91 to 2.90.

Type: Numeric

Algorithm: Factor score computed from the source variables listed below utilizing a model described in ARIC Manual 30.
If VTYPE='V2' and V6DWRC2, V6DSSC2, and V6WFTC2 are missing then GLOBALFS1 is set to missing.
If VTYPE='V4' and V6DWRC4, V6DSSC4, and V6WFTC4 are missing then GLOBALFS1 is set to missing.
If VTYPE='V5NCS' and V6DWRC5, V6WFTC5, V6DSSC5, V6ILRC5, V6ANSC5, V6DSBC5, V6BNTC5, V6LMTC5, V6TMTAC5, and V6TMTBC5 are missing then GLOBALFS1 is set to missing.
If VTYPE='V6NCS' and V6DWRC6, V6WFTC6, V6DSSC6, V6ILRC6, V6ANSC6, V6DSBC6, V6BNTC6, V6LMTC6, V6TMTAC6, and V6TMTBC6 are missing then GLOBALFS1 is set to missing.
If VTYPE='BRAIN' and V6DWRCM, V6WFTCM, V6DSSCM, V6ILRCM, V6ANSCM, V6LMTCM, V6TMTACM, and V6TMTBCM are missing then GLOBALFS1 is set to missing.

Source variable(s): [V6FactorsMplus] VTYPE, V6DWRC2, V6DSSC2, V6WFTC2, V6DWRC4, V6DSSC4, V6WFTC4, V6DWRC5, V6WFTC5, V6DSSC5, V6ILRC5, V6ANSC5, V6DSBC5, V6BNTC5, V6LMTC5, V6TMTAC5, V6TMTBC5, V6DWRC6, V6WFTC6, V6DSSC6, V6ILRC6, V6ANSC6, V6DSBC6, V6BNTC6, V6LMTC6, V6TMTAC6, V6TMTBC6, V6DWRCM, V6WFTCM, V6DSSCM, V6ILRCM, V6ANSCM, V6LMTCM, V6TMTACM, V6TMTBCM

4.2 LANGUAGEFS1 (Language domain factor score from categorized data (ver1))

Description: A language domain factor score estimated from a categorical confirmatory factor analysis of the Word Fluency Test, Animal Naming Score, and Boston Naming Test. A factor score was computed for Visit 5 based on discretized versions of each neurocognitive test. The factor loadings and thresholds from Visit 5 were integrated as fixed parameters into a model that generated factor scores for Visit 6. Factor scores range from -2.03 to 2.11.

Type: Numeric

Algorithm: Factor score computed from the source variables listed below utilizing a model described in ARIC Manual 30.
If VTYPE='V5NCS' and V6WFTC5, V6ANSC5, and V6BNTC5 are missing then LANGUAGEFS1 is set to missing.
If VTYPE='V6NCS' and V6WFTC6, V6ANSC6, and V6BNTC6 are missing then LANGUAGEFS1 is set to missing.

Source variable(s): [V6FactorsMplusLang] VTYPE, V6WFTC5, V6ANSC5, V6BNTC5, V6WFTC6, V6ANSC6, V6BNTC6

4.3 EXECFUNCFS1 (Executive functioning domain factor score from categorized data (ver1))

Description: An executive functioning domain factor score estimated from a categorical confirmatory factor analysis of the Digit Symbol Substitution, Trail Making Test A, and Trail Making Test B. A factor score was computed for Visit 5 based on discretized versions of each neurocognitive test. The factor loadings and thresholds from Visit 5 were integrated as fixed parameters into a model that generated factor scores for Visit 6. Factor scores range from -1.90 to 2.24.

Type: Numeric

Algorithm: Factor score computed from the source variables listed below utilizing a model described in ARIC Manual 30.
If VTYPE='V5NCS' and V6DSSC5, V6TMTAC5, and V6TMTBC5 are missing then EXECFUNCFS1 is set to missing.
If VTYPE='V6NCS' and V6DSSC6, V6TMTAC6, and V6TMTBC6 are missing then EXECFUNCFS1 is set to missing.

Source variable(s): [V6FactorsMplusExecFunc] VTYPE, V6DSSC5, V6TMTAC5, V6TMTBC5, V6DSSC6, V6TMTAC6, V6TMTBC6

4.4 MEMORYFS1 (Memory domain factor score from categorized data (ver1))

Description: A memory domain factor score estimated from a categorical confirmatory factor analysis of the Delayed Word Recall, Incidental Learning, and Logical Memory Test. A factor score was computed for Visit 5 based on discretized versions of each neurocognitive test. The factor loadings and thresholds from Visit 5 were integrated as fixed parameters into a model that generated factor scores for Visit 6. Factor scores range from -2.13 to 2.69.

Type: Numeric

Algorithm: Factor score computed from the source variables listed below utilizing a model described in ARIC Manual 30.
If VTYPE='V5NCS' and V6DWRC5, V6ILRC5, and V6LMTC5 are missing then MEMORYFS1 is set to missing.
If VTYPE='V6NCS' and V6DWRC6, V6ILRC6, and V6LMTC6 are missing then MEMORYFS1 is set to missing.

Source variable(s): [V6FactorsMplusMem] VTYPE, V6DWRC5, V6ILRC5, V6LMTC5, V6DWRC6, V6ILRC6, V6LMTC6

5. REFERENCES

- Benton, A., & Hamsher, K. (1976). *Multilingual Aphasia Examination*. Iowa City, IA: University of Iowa.
- Gross, A., Power, M., Albert, M., Deal, J., Gottesman, R., Griswold, M., . . . Bandeen-Roche, K. (2015). Application of latent variable methods to the study of cognitive decline when tests change over time. *Epidemiology*, 26(6), 878-887.
- Hayden, K., Jones, R., Zimmer, C., Plassman, B., Browndyke, J., Pieper, C., & Welsh-Bohmer, K. (2011). Factor structure of the National Alzheimer's Coordinating Centers uniform dataset neuropsychological battery: an evaluation of invariance between and within groups over time. *Alzheimer Disease and Associated Disorders*, 25(2), 128-137.
- Park, L., Gross, A., McLaren, D., Pa, J., Johnson, J., Mitchell, M., . . . Initiative, A. D. (2012). Confirmatory factor analysis of the ADNI neuropsychological battery. *Brain Imaging and Behavior*, 6(4), 528-539.
- Rawlings, A., Bandeen-Roche, K., Gross, A., Gottesman, R., Coker, L., Penman, A., . . . Mosley, T. (2016). Factor structure of the ARIC-NCS Neuropsychological Battery: An evaluation of invariance across vascular factors and demographic characteristics. *Psychological Assessment*, 28(12), 1674-1683.

- Reitan, R. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271-276.
- Ryan, J., & Lopez, S. (2001). Wechsler adult intelligence scale-III. In W. Dorfman, & M. Hersen, *Understanding psychological assessment. Perspectives on individual differences*. New York, NY: Kluwer Academic/Plenum Publishers.
- Siedlecki, K., Manly, J., Brickman, A., Schupf, N., Tang, M., & Stern, Y. (2010). Do neuropsychological tests have the same meaning in Spanish speakers as they do in English speakers? *Neuropsychology*, 24(3), 402-411.
- Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. San Antonio, Texas: Psychological Corporation.
- Williams, B., Mack, W., & Henderson, V. (1989). Boston naming test in Alzheimer's disease. *Neuropsychologia*, 27(8), 1073-1079.