

WHI Memory Study (WHIMS) Investigator Data Release
Data Preparation Guide
January 2022

1. Introduction

This release consists of a single data set from the WHIMS Epidemiology of Cognitive Health Outcomes (ECHO) study. The data provided are the results of the cognitive battery and information regarding the classification of cognitive status (details provided below) from 2008 through the end of the study (August 31, 2021).

2. Description of WHIMS ECHO

Women who were previously enrolled in the WHIMS Extension were recruited for WHIMS ECHO beginning in 2008. Instead of face-to-face evaluation, participants undergo an annual centralized, validated cognitive telephone interview for tracking changes in cognitive status (Rapp et al., 2012). The interview is comprised of a neuropsychological battery including a global cognitive screener (modified Telephone Interview for Cognitive Status, known as the TICSm) and additional neuropsychological tests.

All participants receive the entire battery (see below). If a woman scores below 31 on the TICSm, the Dementia Questionnaire (DQ) is administered to a friend or family member previously identified (proxy). The DQ is a standardized, validated instrument used to reliably classify dementia when used with cognitive performance measures. A central panel of experts in the diagnosis of mild cognitive impairment syndrome and dementia reviews the results along with the cognitive scoring history and classifies participants as follows: 1) no dementia (no cognitive impairment), 2) mild cognitive impairment (MCI), 3) probable dementia (PD), 4) unable to classify – cognitive impairment, 5) unable to classify – functional impairment, and 6) unable to classify – no cognitive impairment and no DQ.

‘Unable to classify’ designations are made when either key information is missing or factors affecting cognitive or behavioral function (e.g., depression, acute medical illness) are present and preclude a confident classification into one of the three main categories (No Dementia, MCI, Dementia). ‘Unable to classify-cognitive impairment’ indicates that cognitive data are available and were judged to be impaired by but missing data (typically a missing Dementia Questionnaire) or confounding factors are present. Similarly, ‘Unable to classify-functional impairment’ indicates that the DQ reflects significant functional impairment, but because some or all cognitive test data are missing. ‘Unable to classify-no cognitive impairment and no DQ’ is made when cognitive impairment is not clearly evident and the DQ is missing. For analytical purposes this group can be considered ‘Normal’ (No Dementia) since cognitive impairment is a requisite for MCI and Dementia. (NOTE: The California Verbal Learning Test, modified as described below, is not reviewed or considered when adjudicating cases).

3. Data File Setup

The data file is a SAS data set with one or more observations per participant. The identifying variable in each file is Participant ID, 'ID' (referred to as the “common ID” in the WHI documentation). Computed variables have also been included and are described in detail under the next section, contents of the cognitive battery. Refer to the Appendix for further details on the tests and questionnaires.

4. Contents of the ECHO Dataset

Test/Scale	Variable name	Variable description
Not Applicable	ID	Participant ID
Not Applicable	ticsdays	Days from HT Trial randomization to cognitive battery
Not Applicable	dxdays	Days from randomization to final diagnosis
Not Applicable	Final_dx	Classification of participants as follows 1) no dementia (no cognitive impairment) 2) mild cognitive impairment (MCI) 3) probable dementia (PD) 4) unable to classify – cognitive impairment 5) unable to classify – functional impairment 6) unable to classify – no cognitive impairment and no DQ
Not Applicable	Finalclass	Classification of participants with a probable dementia diagnosis into the following classes: 1) Probable Vascular 2) Probable Alzheimer’s 3) Dementia: Mixed Type 4) Dementia: Etiology Unknown 5) Other *Note: classification of PD type stopped in 2013 due to adjudicators feeling that they did not have enough information to determine the classification
Telephone Interview for Cognitive Status-modified (TICS-m)	score_01 to score_16 score_01 score_02 score_03 score_04	Sub-scores for each of the 16 components Name and orientation (0 to 9) Counting backwards (0 to 2) Word list recall (0 to 10) Subtracting 7 from 100 by (0 to 5)

	score_05 to score_08 score_09 to score_10 score_11 to score_12	Responsive naming (0 to 4) Repetition (0 to 2) Naming President and Vice-President of USA (0 to 4)
	score_13 score_14 to score_15 score_16	Finger tapping (0 to 2) Word opposites (0 to 2) Delayed word list recall (0 to 10)
	tics_score	TICS total score (0 to 50)
East Boston Memory Test (EBMT)	ebmt_score	Total score (0 to 12)
Verbal Fluency - Animals	vfa_wordcount	Total number named
Oral Trail Making Test – Part A and B	oraltrail_errorsa, oraltrail_correcta, oraltrail_timea	Total number of errors (0 to 5), total number correct (0 to 25), and total time(secs) for Part A* (≤ 300)
	oraltrail_errorsb, oraltrail_correctb, oraltrail_timeb	Total number of errors (0 to 5), total number correct (0 to 25), and total time(secs) for Part B* (≤ 300)
Digit Span Test	digit_forwardscore	Total score for digits forward (0 to 14)
	digit_backwardscore	Total score for digits backward (0 to 14)
Geriatric Depression Scale (GDS)	gds_score	Total score (0 to 15)
WHI Insomnia Rating Scale (WHIRS)	insomnia	Total score (0 to 20)
TICS Word List Long Delay	ticslong_score	Total score (0 to 10)
EBMT Recall	ebmtrecall_score	Total score (0 to 12)
California Verbal Learning Test (CVLT) - modified	cvlt_correct, cvlt_intrusions, cvlt_repetitions	Total score (0 to 48) Number of intrusions Number of repetitions

* Note that total time (secs) is a character variable; DC indicates that the test was discontinued before completion.

5. Data Conventions

Dates

No actual dates are included in the data files. All dates have been converted to the number of days since WHI randomization. A negative number of days indicate the date occurred before randomization. Likewise, a positive number indicates occurrence after randomization.

Data Edits

At data entry, the built-in features of the study database application prevented entry of most invalid or impossible data values for categorical variables. Broad range checks applied to continuous variables have set out-of-range responses to missing. There still may be values that appear extreme; **it is up to the user to examine all data before proceeding with data analysis.**

Missing Data

Missing data can result from a participant not completing all tests or assessments. Missing values in the data files are represented by a single period (".") for quantitative variables or blank (" ") for character variables.

5. Appending and Merging Data Files

If you wish to expand your data analyses to include WHI Clinical Trial data, you can use the ID variable in the WHIMS data set and the ID variable in the WHI Clinical Trial data set to merge data sets. The WHIMS and WHI Clinical Trial data releases use the same participant ID.

The ECHO data can be appended to the previously released on trial and post-trial/extension data for the classification of cognitive impairment. The variables F39DAYS (from the early release of this data) and TICSDAYS (from the 2015 release/current data) both represent time in days from HT Trial randomization until cognitive testing/adjudicated classification.

For further information about this data release, please contact Katelyn Garcia at kgarcia@wakehealth.edu or Julia Robertson at jurobert@wakehealth.edu.

Appendix

Telephone Interview for Cognitive Status-modified (TICS_m) is a widely-used measure of **global cognitive functioning** modeled after the Mini Mental State Exam (MMSE) (Folstein, 1975; Brandt, 1988; Welsh, 1993). The TICS_m is a 16-question test (range of scores 0 to 50) with items assessing subject's name, telephone number, date, counting backwards, word list recall, subtraction, responsive naming, repetition, President's name, Vice President's name, finger tapping, word opposites and a delayed word list recall.

East Boston Memory Test (EBMT) measures **verbal memory** (Albert, 1991). Participants are read a short paragraph consisting of 12 distinct elements and immediately asked to recall as many elements as possible (immediate recall) and again 15 minutes later (delayed recall).

Oral Trail Making Test (OTMT) is a modified version of the original TMT (Reitan, 1992), a widely used and well-validated measure of **attention** (Part A) and **executive function** (Part B). For Part A participants are asked to count from 1 to 25 as fast as they can. Part B requires them to recite an alternating numeric and alphanumeric sequence (1-A-2-B-3-C...13) as quickly as possible. The time to complete each task in seconds is the score.

Verbal Fluency-Animals (VF-A) measures **verbal fluency** (Benton, 1968). This task requires participants to spontaneously name as many animals as possible during 1 minute. The total number of unique words is the score.

Digit Span Test (DST) measures **working memory** (Wechsler, 1981). This task requires the participant to repeat sequentially series of single digit numbers of increasing length presented orally, first as presented (Digits Forward) and subsequently in reverse order (Digits Backwards). The score is the longest span of digits recalled. The sum of Digits Forward and Digits Backward (DST-Total) will be used.

Geriatric Depression Scale-Short Form (GDS_SF) measures **depression** (Yesavage, 1988; Burke, 1991). This 15-item (Y/N) questionnaire, which can be administered orally, has good psychometric properties, and has good normative data (Osborn, 2002; De Craen, 2003).

Women's Health Initiative Insomnia Rating Scale (WHIIRS) measures **sleep disturbance**, which is related to cognitive function and hormonal variations (Levine, 2003). This 5-item self-report has excellent reliability, construct validity and is sensitive to change over time.

California Verbal Learning Test (CVLT) measures **verbal learning** and **verbal memory**. It was modified to consist of 3 learning trials of the same 16-item word list with immediate recall of as many words as possible after each trial (Immediate Recall) to reduce administration time and participant burden.

Dementia Questionnaire (DQ) is a semi-structured interview designed for a knowledgeable proxy to provide information needed to make a dementia and MCI diagnoses and to identify causes of cognitive impairment (Kawas et al., 1994). It covers six domains: memory and cognition, verbal expression, daily functioning, recognition of problems/insight, other medical and psychiatric problems, and medical contacts. Knowledgeable friends or family members also estimate the years of symptom onset.