Memento
User manual

Part 5. Datasets and statistical considerations

Version Septembre 2021

TABLE OF CONTENTS

Electronic Case Report Form datasets

Datastet	Description	Commentary
ACC	Informant Questionnaire	Study-partner demographics and relationship with the patient
AD8	AD8 Questionnaire	Informant-reported cognitive complaint based on phone interview
ADL	Activities of Daily Living (ADL)	Basic activities of everyday life
AE	Adverse Events	
AIDE	Help at home	
ATCDP1	Medical history	Cardio-vascular and others (closed list)
ATCDP2	Medical history	Eye and menopause
BREF	Frontal Assessment Battery	Part of the Neuropsychological Battery
CDR	Clinical Dementia Rating scale	
CESD	Informant Depression scale (CES-D)	Presence of depressive symptoms in the study-partner
CM	Medications	
COHEN	Cohen et al. scale	Level of stress perceived by the study-partner
COINF	Informant contact sheet	
COMED	Physician contact sheet	
DC	Death	
DM	Demography	
DMS	Delayed Matching Sample (DMS-48)	Memory test, part of the Neuropsychological Battery
DO	Name Object 80	Language test, part of the Neuropsychological Battery
EAV	Subjective complaints	Completed by the patient themself
ELP	Fronto-temporal behavioural scale (Lebert and Pasquier)	Filled only if demented
EMP	Numeric and visuo-spatial Span	Short-term memory/executive test, part of the Neuropsychological Battery. Visuo-spatial part was optional.
EQ5D	EuroQol EQ-5D questionnaire	Participant's quality of life
EV	Life events	
EXAM	Clinical examination	
EXAMC	Clinical examination (morphometry, blood pressure, heart rate)	
EXC	Examination planning	
FLU	Verbal fluencies	Language/executive test, part of the Neuropsychological Battery. 2 min letter P and Animals

FMH	Familial medical history	
HISM	History of the disease	Current disorders/disease
HV	Tobacco and alcohol consumption	
IADL	Instrumental Activities of Daily Living (IADL)	everyday activities more complex than ADLs
INCL	Inclusion	
IPAQ	International Physical Activity Questionnaire	
IRM	MRI (visual quality check, hippocampal and white matter lesions visual scales)	Scheltens and Fazekas scales
LAB	Standard blood sample labs	
LEIP	LEIPAD questionnaire	study-partner's quality of life as reported by the study-partner themself
LOI	Leisure activities	
MCK	Dementia with Lewy bodies (McKeith criteria)	Filled only if demented
MCL	Lewy body disease evaluation questionnaire (part 1)	Optional
MCLQ	Lewy body disease evaluation questionnaire (part 2)	Optional
MH	Medical history	
MMSE	Mini Mental State Examination (MMSE)	
MNA	Mini Nutritional Assessment (MNA)	
NPIC	Neuropsychiatric Inventory Questionnaire (execution)	
NPICA	NPI (informant part)	
NPICQ	NPI (clinician part)	
NPIR	NPI (phone version)	
OCU	Eye monitoring	
PECP	Medical and paramedical care	
PRAX	Praxis (Mahieux et al)	Part of the Neuropsychological Battery
QTEL	Phone interview	When visit was not done on clinical site
RES	Neuropsychological Tests battery results at inclusion	
REY	Rey figure	Visuospatial/executive test, Part of the Neuropsychological Battery. At 30min was optional
RLRI_IMD	Free and Cued Selective Reminding Test (part 1)	Memory test, Part of the Neuropsychological Battery
RLRI_SC	Free and Cued Selective Reminding Test (part 2)	Memory test, Part of the Neuropsychological Battery
RSO	Social network	
SCHAC	Hachinski Ischaemic score	Filled only if demented
SERO	Serology	If clinically relevant
SOCIO	Socio-economic characteristics	

SPPB	Short Physical Performance Battery (SPPB)	
STVIS	Summary of visit attendance	Include a time-dependent site variable for participants who moved
SU	Subjective difficulties	Self-reported cognitive complaint
SYNTH	Summary of the medical examination	
TMT	Trail Making Test (TMT)	Executive test, Part of the Neuropsychological Battery
TNP	Neuropsychological Tests execution	
ZARIT	Zarit Informant burden inventory	Caregiver burden as reported by the study-partner

Imaging Datasets

Datastet	Description	Commentary
AMYLOIDE_T1_T2_LO	Amyloïd PET (Amyging+INSIGHT, longitudinal pipeline)	For participants enrolled in MEMENTO-Amyging ancillary study or INSIGHT study. Global and local SUVR + cut off points to define "positivity" are provided.
AMYLOIDE_T1_T2_TR	Amyloïd PET (Amyging+INSIGHT, transversal pipeline)	For participants enrolled in MEMENTO-Amyging ancillary study or INSIGHT study. Global and local SUVR + cut off points to define "positivity" are provided.
CATI_MONITO	MRI, FDG and amyloid PET execution (from CATI)	Information on date, machines and quality checks.
DIFFUSION	MRI Diffusion (M0)	Optional. Analysis of deep white matter bundles
FMRI	Functional MRI (M0 Resting state)	Optional. Analysis of functional networks integrity
FOLDS_M0_M24_TR	Cortical folds opening (transversal pipeline)	Morphologist pipeline
FREEAREA_M0_M24_LO	FreeSurfer cortical surfaces (longitudinal pipeline)	FreeSurfer 5.3
FREEVOL_M0_M24_LO	FreeSurfer volumes (longitudinal pipeline)	FreeSurfer 5.3
HSB_M0_M24_TR	White matter hyperintensities (transversal pipeline)	WHASA pipeline
SPMVOL_M0_M24_TR	SPM Global brain volumes (transversal pipeline)	SPM12
TEP_FDG_M0_M24	18-FDG PET	18-FDG-PET was optional. Global and local SUVR are provided.

THICK_M0_M24_LO	Cortical thickness (longitudinal pipeline)	FreeSurfer 5.3
VOL_HIPP_AMYG_TR_LO	Hippocampal and amygdala volumes (transversal pipeline)	SACHA pipeline for transversal M0. FreeSurfer 5.3 longitudinal pipeline for M0 and M24

Biological samples datasets

Datastet	Description	Commentary
APOE4	APOE genotype	
BIOBANQUE	Biobank storage	Date and time of biobank samples. Samples stored in biobank for subsequent processing might differ of those in the LAB dataset.
CSF	CSF biomarkers	Based on INNOTEST Fujirebio kits
GENE	Alzheimer specific SNPs	Based on Lambert et al. Nat Genet 2013, list of genes
INFLAM_M0	Inflammatory markers M0	From blood and CSF samples
QUANTERIX	Total Tau and P-Tau measures in CSF and Plasma	Quanterix kits

Derived datasets

The following datasets are reshaped versions of existing datatsets (CM, MH, NPI,…), easier to manage. They were designed for analysis purpose and have to be use in priority.

Datastet	Description	Commentary
CM_FOLLOW	Medication exposure by visit	
CM_FOLLOW_RESUM	HBP, diabetes, dyslipidemia medications exposure by visit	
CM_RWK	Medications (cleaned)	
LAB_RWK	Standard blood sample labs (cleaned)	A clean version of the LAB dataset. All measures are provided in international standard unit. Extreme values were removed. To be use in priority for analysis.
LAB_TRANSP	Standard blood sample labs (wide format)	A transposed version of the LAB_RWK dataset.
MH_RWK	Medical history (cleaned)	Gather information from ATCDP1 and MH files.
NPI_RWK	NPI (cleaned)	
PHONESUM	Phone interview (summary)	

Special datasets

The following datasets require a specific attention to the analyst.

DEMENCE

All dementia cases were reviewed by an Endpoint Review Committee (ERC). The "Demence" dataset includes the information provided by the ERC about the presence of dementia (DECISION_DEM), the date of onset (DEMENCE_DAT), the respective Memento visit in the clinical site (DEMENCE_VIS), and two hypotheses of etiology (ETIOl_1, ETIOL_2).

Whenever information is needed about the presence or phenotyping of dementia, this dataset must be consulted as it is validated by the ERC.

Information on the SYNTH dataset comes from the clinical center and might differ.

SEUILS_MCI_VIS

This dataset provides summary information on how the participant performed the neuropsychological test battery. Whenever possible, sex-age-education norms in general population were considered. A threshold of 1.5SD was then used to define an impairment.

- Memory tests were: DMS-48, Free and Cued Selective Reminding Test, and Rey 3min;
- Executive functions tests: TMT A and B, Frontal Assessment Battery and Verbal Fluencies
- Language tests: Verbal Fluencies, Naming Object 80

Details are also given for TMT, Praxis, Digit Span and FAB.

The final MCI classification (MCI_VIS) follows the Petersen 2004 criteria:

- Non MCI: no domain impaired
- Pure aMCI: memory impaired solely
- Pure naMCI: a single non memory domain impaired
- Multidomain aMCI: both memory and at least 1 non memory domain impaired
- Multidomain naMCI: at least two non-memory domains impaired, memory preserved

An extended version is provided (MCI2_VIS), where the non MCI category is divided in non MCI or 1SD impairment (considered sometimes as "early MCI").

The analyst must pay attention that this MCI definition is based on neuropsychological test battery solely and might differ from the clinical evaluation in the CDR.

OVERVIEW

The OVERVIEW dataset must be considered as a key table for all analysis.

This dataset aims at giving a glance of the participants' milestone in the study: follow-up, last news,

events, imaging and CSF availability.

This dataset summarizes information from several datasets in one place and might save a lot of time to the analyst. The dementia information comes from de DEMENCE dataset.

In addition, several variables were designed in a survival analysis perspective:

- Date of consent (INCCONSDAT_D) : the entry time in the study
- EVT: occurrence of dementia, death (free of dementia) or censoring
- EVT_PY: delay in years between date of consent and EVT_DAT
- EVT_DEM: a version of EVT where death is considered as censoring.
- EVT_DEM_PY: obvious…
- INSTITUT_DAT: date of institutionalization, if so
- DERNOUV_DAT: date of last information
- DERNOUV_VIVANT: date of last information alive
- LAST_VIS: last Memento visit

PATSUM

The PATSUM dataset is basically the "baseline" data, i.e., a collection of the main characteristics of the participants at entry in the study.

Information come from several tables (the variables names are unchanged), and few new variables were created. In most cases, their use is obvious, but some deserve more information.

- CM_HTA, CM_DIAB, CM_DYSLIP: based on CM dataset. Coded "Yes" if the participant was exposed to respectively high blood pressure, diabetes or Dyslipidemia drugs at their inclusion.
- PAS, PAD, F_CARD: mean of available (3,2, or 1) measurement of systolic BP, diastolic BP and heart rate.
- HTA_140_90: Coded Yes if either SBP≥140 or DBP≥90
- HTA_140_90_TRT: based on HTA_140_90, but coded Yes if the participant was exposed to a HBP drug.
- HTA_140_90_PEC: a 4 categories variable, depending on the concomitance of HBP and treatment.
- ATCD_CV: coded Yes if the participant experienced at least one of the following event: stroke, angina pectoris, arteriosclerosis obliterans, cardiac failure, myocardial infarction, coronary disease.
- TMTA_TAUX, TMTB_TAUX: time for TMT realisation divide by the number of good moves
- ASS_MIN_J, IPAQ_MET_MIN_W, IPAQ_TOT_CLA: computed following the Guidelines for Data Processing of the IPAQ
- RESTRICTION_CLA: Presence of restriction to IADL, based on Phone, transportation, Medication and Finance items.
- MCI and MCI2: see SEUILS_MCI_VIS dataset description

FOLLOWSUM

The FOLLOWSUM is roughly the longitudinal? version of PATSUM. Information on death and dementia comes from OVERVIEW and DEMENCE datasets.

Specific variables were added:

- CEN_ANOM_VIS: the center where the participant actually went for the clinical visit (i.e., new center for those who moved).
- ADL_RES: restriction to ADL

Statistical considerations

Statistical software

The Memento database was initially created for a SAS framework (labels, formats, missing data…). However, we provide a RData file and txt files (tab delimited).

Missing data

Note that for numeric and date variables, you might find some strange "C", "A", "D", "NC", "NA", "ND", … values. These values are basically missing data. Attention must be paid on how the statistical software deals with these values.

Despite the efforts made to get as comprehensive as possible information, some dates could be partially missing (day, month, or both). This is more frequent in the CM and MH datasets, since they include information prior to inclusion in Memento. Most of time, 2 variables are provided: a text version and a date version. The latter was imputed, if necessary, as follows: missing day => 15th of the month, missing day + month: 1st of july. If no text variable is provide, all dates are exact and not imputed (imaging data for instance).

Time

The "time 0" of the cohort is the day of the signature of the consent (var INCCONSDAT_D). Age, duration, follow-up, delays, etc… must refer to this date (for survival analysis for instance). The OVERVIEW dataset might help the analyst for this point.

Delays are usually computed as (date of interest – INCCONSDAT_D)/365.25 for years, and (date of interest – INCCONSDAT_D)/30.4375 for months.

Correspondence visit number / months	
1	baseline
2	6 months
3	12 months
4	18 months
5	24 months
6	30 months
7	36 months
8	42 months
9	48 months
10	54 months
11	60 months

Prevalent cases

After review by the Endpoint Committee Review, 18 participants were classified demented at the time of their inclusion. These individuals were followed in the clinical center and their data collected following the same process as the other participants. We recommend their exclusion of the analytic sample, and they may be systematically excluded from analyses focusing on incident dementia.

We encourage the project team to be familiar of the analysis and reporting good pratices. A good starting point could be:

the STROBE recommendations (https://www.equator-network.org/reporting-guidelines/strobe/)

The TRIPOD recommendations (https://www.equator-network.org/reporting-guidelines/tripod-statement/) for prediction models

The MELODEM princeps paper (Weuve J, Alzheimer's & dementia, 2015)