ASAP CRN Data Dictionary

Overview

This document provides definitions for the variables harmonized across the ASAP CRN based on our Common Data Elements (CDE). These metadata enable the harmonization of the data submitted by multiple teams. The metadata consists of five tables, which are described below.

Tables

STUDY. The `STUDY` table provides comprehensive details about a research project, including its objectives, the team involved, funding sources, related publications, and specifics about the samples used. This top level information is the context under which the post-mortem brain sample RNA sequences were derived.

PROTOCOL. The `PROTOCOL` table describes the methodologies employed in the study, detailing how samples were collected, processed, and analyzed. This ensures transparency and reproducibility in the research.

SUBJECT. The `SUBJECT` table provides information about individual subjects (or donors) from whom samples were collected. Individual Subjects can contribute multiple Samples. This table details the most relevant demographic and phenotypic aspects of the subject.

SAMPLE. The 'SAMPLE' table has the specific details for each individual brain sample, which subject it is derived from, and the name of the raw data artifact.

CLINPATH . The Clinical Pathology (`CLINPATH`) table contains detailed clinical and pathological data (from the donating brain bank) for each sample.

DATA . The `DATA` table details the data files associated with each `SAMPLE`.

Table Key Values

SUBJECT and SAMPLE tables are connected via the key *subject_id* field.

SUBJECT and CLINPATH tables are connected via the key value *subject_id* and *source_subject_id* fields

SAMPLE and DATA tables are connected via the key value `sample_id` fields

ASAP Assigned Table Key Values

Note that there are also additional ASAP Key Values assigned centrally. These unique keys will link each table in a dataset, and multiple datasets together.   The crucial purpose of these is to ensure that samples from any individual subject across teams is identified.

STUDY	ASAP_dataset_id
SUBJECT	ASAP_subject_id
SAMPLE	ASAP_sample_id
SAMPLE	ASAP_dataset_id
PROTOCOL	ASAP_dataset_id
CLINPATH	ASAP_subject_id

V1 to v2 changes

Overall the logic of *study*, *subject*, and *sample* -level metadata was clarified, and to simplify the generation and maintenance of the ASAP Assigned fields detailed above. The main evolution of the CDE is to break a table of files and raw data details out from the SAMPLE table into a DATA table. Additionally the CLINPATH table shared keys were changed to map to SUBJECT instead of SAMPLE making it a truly *sample* level table, enabling subject fields NOT directly related to the clinical pathology to be placed in the SUBJECT table.
A detailed summary of the changes can be found in ASAP CDE ALL google document in the 'v1-v2 Changes' sheet.

Data Dictionaries:

STUDY

Table	Column Name	Data Type	Required	Description	Enum Values
STUDY	ASAP_team_name	Enum	Required	ASAP Team Name: Name of the ASAP CRN Team. i.e. "TEAM-LEE","TEAM-HAFLER","TEAM-HARDY",etc.	["TEAM-LEE","TEAM-HAFLER","TEAM-HARDY", "TEAM-JAKOBSSON", "TEAM-SCHERZER", "TEAM-SULZER", "TEAM-VOET", "TEAM-WOOD"]
STUDY	ASAP_lab_name	String	Required	Lab Name. : Lab name that is submitting dataset for the ASAP CRN Team.	
STUDY	project_name	String	Required	Project Name: A Title of the overall project. A unique and clear title of the Team's project.	
STUDY	team_dataset_id	String	Required	The "project_name" is often too verbose for practical use, so adding a short "ID" will be helpful for coding team components of datasets (i.e. ASAP_dataset_id) Team level dataset name. A one or two word descriptor joined by "_" to distinguish this dataset from other datasets this team might submit.	
STUDY	project_dataset	String	Required	Dataset Name: A unique name is required for each data submission, as one project can generate and submit multiple datasets.	
STUDY	project_description	String	Required	Project Description: Brief description of the goals and objectives of the project.	
STUDY	PI_full_name	String	Required	Principal Investigator Name.: Full name of the lab Principal Investigator. Format: First name, Initial, Last name	
STUDY	PI_email	String	Required	Principal Investigator E-mail.: E-mail address of the Principle Investigator of the Lab submitting the data.	
STUDY	contributor_names	String	Required	Contributor(s).: Names of all contributors to the dataset. Format: First name, Initial, Last name ; First name, Initial, Last name ; ...	

STUDY	submitter_name	String	Required	Data Submitter Name.: Full name of individual submitting data.	
STUDY	submitter_email	String	Required	Data Submitter E-Mail.: Email address of individual submitting data.	
STUDY	ASAP_grant_id	String	Required	ASAP Grant ID.: ID of the ASAP grant funding the study. e.g. ASAP-000xxx	
STUDY	other_funding_source	String	Required	Other Funding.: Details of other funding contributing to data generation.	
STUDY	publication_DOI	String	Required	bioRxiv DOI.: Digital Object Identifer for pre-print(s). NA if not available.	
STUDY	publication_PMID	String	Required	PubMed ID.: Unique identifier for journal articles/abstracts in PubMed. NA if not available.	
STUDY	number_of_brain_samples	Integer	Required	Number of Brain Samples.: Total number of brain tissue samples in the dataset.	
STUDY	brain_regions	String	Required	Brain Regions.: List of all anatomical origins of brain samples. This should be as specific as possible for this field.	
STUDY	types_of_samples	String	Required	Sample Type. : Description of samples. Currently, this is ONLY: late-stage PD and control postmortem brains	
STUDY	PI_ORCHID	String	Optional	PI ORCID.: Principle investigator	
STUDY	PI_google_scholar_id	String	Optional	PI Google Scholar ID.:	
STUDY	DUA_version	String	Required	DUA Version. : ASAP Access & Use Policy	
STUDY	preprocessing_references	String	Optional	Preprocessing References.: Github repository link or similar code reference plus electronic lab notes detailing all data generation steps and quality control summaries. Analysis scripts for additional processing will also create valuable context.	
STUDY	metadata_version_date	String	Optional	Metadata Version and Date. : Metadata Version and Date	
STUDY	alternate_dataset_id	String	Optional	Alternate dataset identifier	

PROTOCOL

Table	Column Name	Data Type	Required	Description	Enum Values
PROTOCOL	sample_collection_summary	String	Required	Sample Collection Summary.: Short summary detailing how samples were collected.	
PROTOCOL	cell_extraction_summary	String	Required	Cell Extraction Summary.: Short summary detailing the protocols used for cell extraction.	
PROTOCOL	lib_prep_summary	String	Required	Library Preparation Summary.: Short summary detailing how the library was prepared.	
PROTOCOL	data_processing_summary	String	Required	Data Processing Summary.: Short summary of how data has been processed prior to fastq generation.	
PROTOCOL	github_url	String	Required	Github Repository Link.: URL for lab github repo conatining the code and scripts for processing data in this dataset.	
PROTOCOL	protocols_io_DOI	String	Required	Protocols.io DOI.: Digital Object Identifier (DOI) for protocols.io	
PROTOCOL	other_reference	String	Optional	Other Reference.: Link to technology/platform protocol(s) if not available in protocols.io	

SUBJECT

Table	Column Name	Data Type	Required	Description	Enum Values
SUBJECT	subject_id	String	Required	The Subject ID. This will be a unique ID for the subject. Please submit the subject ID used internally. In the harmonized dataset this may be modified by the ASAP team to prevent naming collisions.	
SUBJECT	source_subject_id	String	Required	Source Subject ID.: Subject ID issued by providing brain bank.	
SUBJECT	AMPPD_id	String	Required	AMP-PD ID.: The AMP-PD ID if subhect has samples in AMP-PD. NULL if not in AMP-PD	
SUBJECT	GP2_id	String	Required	GP2 ID if subject has samples in GP2. NULL if not in GP2	
SUBJECT	biobank_name	String	Required	Biobank Name.: Name of the biobank from which the samples were sourced.	
SUBJECT	organism	Enum	Required	Organism.: Organism of the subject. Currently this is ONLY: "Human"	["Human", "Mouse","Dog","Fly","Other"]
SUBJECT	sex	Enum	Required	Sex.: Genetically derived sex.	["Male", "Female","Intersex","Unnown"]
SUBJECT	age_at_collection	Float	Required	Age at Collection.: Numerical age of subject at sample collection. Unit of years in decimal format.	
SUBJECT	race	Enum	Required	Race.: Reported race of subject. NINDS CDE for race. NOT genetic race.	["American Indian or Alaska Native", "Asian", "White", "Black or African American", "Multi-racial", "Native Hawaiian or Other Pacific Islander", "Other", "Unknown", "Not Reported"]

SUBJECT	ethnicity	Enum	Required	Ethnicity.: Reported ethnicity of subject. NINDS CDE for ethnicity. NOT genetic race.	["Hispanic or Latino", "Not Hispanic or Latino", "Unknown", "Not Reported"]
SUBJECT	family_history	Enum	Required	Family History.: Answer to: Did your parents, full-siblings, or children have Parkinson's disease?	["Yes","No","Unknown", "Not Reported"]
SUBJECT	last_diagnosis	String	Required	Last Diagnosis.: Most recent "clinical" diagnosis.	
SUBJECT	age_at_onset	Integer	Required	Age at Onset.: Age in years at which any sympton was manifest, Between 0 and 120	(y>=0) & (y<=120)
SUBJECT	age_at_diagnosis	Integer	Required	Age at diagnosis: Age in years at which clinical diagnosis was made. Between 0 and 120	(y>=0) & (y<=120)
SUBJECT	first_motor_symptom	Integer	Required	First Motor Symptom.: Age in years at which first motor symptoms manifest. Between 0 and 120	(y>=0) & (y<=120)
SUBJECT	hx_dementia_mci	Enum	Required	Dementia MCI.: Answer to: Have you ever been diagnosed with Dementia or MCI?	["Yes", "No"]
SUBJECT	hx_melanoma	Enum	Required	Melanoma.: Answer to: Have you ever been diagnosed with Melanoma?	["Yes", "No"]
SUBJECT	education_level	Enum	Required	Education Level.: Derived value for level of education.	["High School", "High School/GED", "Some college without degree", "Associate degree college", "Bachelor's degree", "Master's degree", "Professional or doctoral degree", "Refuse", "Other"]
SUBJECT	smoking_status	Enum	Required	Smoking Status.: Assessment of "substantial" smoking.	["Current smoker", "Former smoker", "Never", "Unknown"]
SUBJECT	smoking_years	Float	Optional	Smoking Years.: Time in years of "substantial" smoking.	
SUBJECT	APOE_e4_status	Enum	Required	APOE e4 Status.: APOE e4 alelle phenotype.	["22", "23","24","33", "34", "44", "Unknown"]

Table	Column Name	Data Type	Required	Description	Enum Values
SUBJECT	cognitive_status	Enum	Required	Cognitive Status.: Derived cognitive status from general cognitive screening test (i.e., MMSE, MoCA, CDR, etc.).	["Normal", "MCI", "Dementia"]
SUBJECT	time_from_baseline	String	Required	Time from baseline visit in days, baseline is when the participants were enrolled: Time from baseline visit in days, baseline is when the participants were enrolled	
SUBJECT	primary_diagnosis	Enum	Required	Primary Diagnosis.: The diagnosis of the subject. Because subjucts could have multple diagnoses, this is the PRIMARY diagnosis.	["Healthy Control", "Idiopathic PD", "Alzheimer's disease", "Frontotemporal dementia", "Corticobasal syndrome", "Dementia with Lewy bodies", "Dopa-responsive dystonia", "Essential tremor", "Hemiparkinson/hemiatrophy syndrome", "Juvenile autosomal recessive parkinsonism", "Motor neuron disease with parkinsonism", "Multiple system atrophy", "Neuroleptic-induced parkinsonism", "Normal pressure hydrocephalus", "Progressive supranuclear palsy", "Psychogenic parkinsonism", "Vascular parkinsonism", "No PD nor other neurological disorder", "Spinocerebellar Ataxia (SCA)", "Prodromal non-motor PD", "Prodromal motor PD", "Other neurological disorder"]
SUBJECT	primary_diagnosis_text	string	Optional	Primary Diagnosis Text.: If the primary_diagnosis is "Other neurological disorders" diagnosis must be specified here.	

SAMPLE

Table	Column Name	Data Type	Required	Description	Enum Values

SAMPLE	sample_id	String	Required	Sample ID.: ID of each individual sample used in the study (e.g. cortex of anterior cingulate gyrus); if from the same brain several regions are included in the study, each region (i.e. each sample) has different ID and listed as a separate entry	
SAMPLE	subject_id	String	Required	Subject ID.: Can be multiple samples from each Subject	
SAMPLE	source_sample_id	String	Required	Source Sample ID.: ID provided by Brain bank or other source	
SAMPLE	replicate	String	Required	Replicate.: Indicator for replicates of the sample. Replicates could be biological or tecnical. Indicate biological replicates as capitalized "Rep" + replicated number (i.e. Rep1, Rep2, etc) and technical replicates with lowercase "rep" + replicate number (i.e., rep1, rep2 etc.)	
SAMPLE	replicate_count	Integer	Required	Replicate Count.: Number of replicates of this sample. Typically this will be 1, but if replicates have been collected indicate how many total.	
SAMPLE	repeated_sample	Integer	Required	Repeated Sample.: Repeated Sample indicator. This will be treated as a boolean. i.e. enter 0 unless sample is a replicate.	
SAMPLE	batch	String	Required	Batch.: Label to enumerate the batch. This will probablby numerica, but could be coded as letters.	
SAMPLE	tissue	String	Required	Tissue.: Type of tissue. e.g. "Brain" or other organ. For this database it should always be "Brain".	
SAMPLE	brain_region	String	Required	Brain Region.: The anatomical origin of brain sample.	
SAMPLE	hemisphere	Enum	Required	Side of the hemisphere (i.e. right or left) from which the sample is acquired: Side of the hemisphere (i.e. right or left) from which the sample is acquired	["Right","Left","Unknown"]
SAMPLE	region_level_1	Enum	Required	Region Level 1.: General (Level 1) anatomical location.	["Frontal lobe", "Temporal lobe", "Parietal lobe", "Occipital lobe", "Cingulate gyrus", "Striatum", "Thalamus", "Midbrain", "Pons",

					"Medulla", "Cerebellum", "Intestine", "Unknown"]
SAMPLE	region_level_2	Enum	Required	Region Level 2.: More specific (Level 2) anatomical subdivision. Note that "CA1-CA4" indicates Hippocampus origin.	["Prefrontal cortex","Superior frontal gyrus","Middle frontal gyrus","Inferior frontal gyrus","Superior temporal gyrus","Middle temporal gyrus","Inferior temporal gyrus","Fusiform gyrus","Transentorhinal region","Entorinal region","Subiculum","CA1-CA4","Amygdala","Periamygdala cortex","Anterior cingulate gyrus","Posterior cingulate gyrus","Superior parietal lobule","Inferior parietal lobule","Parastriate cortex","Peristriate cortex","Striate cortex","Insular cortex","Caudate nucleus","Putamen","Globus pallidus","Thalamus","Subthalamic nucleus","Substantia nigra","Pontine tegmentum","Pontine base","Medulla tegmentum","Medulla base","Cerebellar vermis","Cerebellar hemisphere","Dentate nucleus","Right colon","Left colon","Unknown"]
SAMPLE	region_level_3	Enum	Required	Region Level 3.: Specification if the sample is from the grey or white matter.	["Grey matter", "White matter", "Grey and white matter", "Unknown"]
SAMPLE	RIN	Float	Required	RIN.: RNA integrity number (RIN) is an algorithm for assigning integrity values to RNA measurements. RIN on sample received (performed by team/lab/3rd party)?	
SAMPLE	source_RIN	Float	Required	Source RIN.: RIN provided by Tissue/Brain/Biobank, if applicable.	
SAMPLE	molecular_source	Enum	Required	Molecular Source.: The molecular source of the data could be DNA, RNA or protein, for example. Initially this will be RNA.	["Total RNA", "PolyA RNA", "Cytoplasmic RNA", "Nuclear RNA", "Genomic DNA", "Protein", "Other"]

SAMPLE	`input_cell_count`	Integer	Required	Input Cell Count.: Total count of cells (single cell or single nuclei) that were originally isolated and processed for capture. I.e. cell suspension concentration * volume injected.	
SAMPLE	`assay`	Enum	Required	Assay.: Assay performed and Version.	["v3.1 - Dual Index","v3.1 - Single Index","v2 - Dual Index", "v1", "Xenium In Situ","v1.1 - Single Index", "v2"]
SAMPLE	`sequencing_end`	Enum	Required	Sequencing End.: Specification sequencing type: Single End vs. Paired-End.	["Single", "Paired-end"]
SAMPLE	`sequencing_length`	Enum	Required	Sequencing Length.: Number of base pairs sequenced.	["25", "50", "100", "150"]
SAMPLE	`sequencing_instrument`	Enum	Required	Sequencing Instrument.: The instrument used to perform sequencing.	["Illumina NovaSeq 6000", "Ilumina HiSeq 2500", "Illumina HiSeq 3000", "Illumina HiSeq 4000", "Illumina NextSeq 550", "Illumina HiSeq X Ten", "Illumina HiSeq X Five", "Illumina MiSeq", "on Torrent PGM", "Ion Torrent Proton", "Ion Torrent S5", "Ion Torrent S5 XL", "PacBio RS", "PacBio RS II", "Other"]
SAMPLE	`organism_ontology_term_id`	Enum	Required	Organism Ontology Term ID.: NCBITaxon	["NCBITaxon:9606"]
SAMPLE	`development_stage_ontology_term_id`	String	Required	Development Stage Ontology Term.: HsapDv if human. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/hsapdv	

SAMPLE	sex_ontology_term_id	Enum	Required	Sex Ontology Term ID.: PATO:0000384 for male, PATO:0000383 for female, or unknown if unavailable	["PATO:0000384 (male)", "PATO:0000383 (female)", "Unknown"]
SAMPLE	self_reported_ethnicity_ontology_term_id	String	Required	Self Reported Ethnicity Ontology.: HANCESTRO use multiethnic if more than one ethnicity is reported. If human and information unavailable, use unknown. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/hancestro	
SAMPLE	disease_ontology_term_id	String	Required	Disease Ontology Term ID.: MONDO or PATO:0000461 for 'normal'. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/mondo https://github.com/chanzuckerberg/single-cell-curation/blob/main/schema/3.0.0/schema.md#suspension_type	
SAMPLE	tissue_ontology_term_id	String	Required	Tissue Ontology Term ID.: UBERON. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/uberon	
SAMPLE	cell_type_ontology_term_id	String	Required	Cell Type Ontology Term ID.: CL. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/cl	
SAMPLE	assay_ontology_term_id	String	Required	Assay Ontology Term ID.: EFO. Included for compatibilty with CELLxGENE. https://www.ebi.ac.uk/ols/ontologies/efo	
SAMPLE	suspension_type	String	Required	Suspension Type.: Cell, nucleus, or na, as corresponding to assay. Included for compatibilty with CELLxGENE. Use https://github.com/chanzuckerberg/single-cell-curation/blob/main/schema/3.0.0/schema.md#suspension_type defined in the data schema for additional guidance.	
SAMPLE	DV200	Float	Optional	DV200.: Value is the percentage of RNA fragments larger than 200 nucleotides.	
SAMPLE	pm_PH	Float	Optional	Post-mortem PH.: PH measured in the tissue. PH is designated as a number ranging between 0 and 14.	(0-14)
SAMPLE	donor_id	String	Optional	Donor ID.: free-text identifier that distinguishes the unique individual that data were derived from. Included for compatibilty with CELLxGENE.	

DATA

Table	Column Name	Data Type	Required	Description	Enum Values
DATA	sample_id	String	Required	Sample ID.: ID of each individual sample used in the study (e.g. cortex of anterior cingulate gyrus); if from the same brain several regions are included in the study, each region (i.e. each sample) has different ID and listed as a separate entry	
DATA	replicate	String	Required	Replicate.: Indicator for replicates of the sample. Replicates could be biological or tecnical. Indicate biological replicates as capitalized "Rep" + replicated number (i.e. Rep1, Rep2, etc) and technical replicates with lowercase "rep" + replicate number (i.e., rep1, rep2 etc.)	
DATA	replicate_count	Integer	Required	Replicate Count.: Number of replicates of this sample. Typically this will be 1, but if replicates have been collected indicate how many total.	
DATA	repeated_sample	Integer	Required	Repeated Sample.: Repeated Sample indicator. This will be treated as a boolean. i.e. enter 0 unless sample is a replicate.	
DATA	batch	String	Required	Batch.: Label to enumerate the batch. This will probablby numerica, but could be coded as letters.	
DATA	file_type	Enum	Required	File Type.: Type of file uploaded for this sample. We are asking for "fastq" for raw data transfer.	["fastq","Per sample raw file", "Per sample processed file", "Combined analysis files", "annData", "vHDF", "plink2", "VCF", "csv"]
DATA	file_name	String	Required	File Name.: Full name of file containing the sample data. For	

				paired-end data (e.g. `sequencing_end` = "single") the ROOT of the file can be specified, and the others will be inferred.	
DATA	`file_description`	String	Required	File Description.: Short description of the file.	
DATA	`file_MD5`	String	Required	File MD5.: Checksum hash of the uploaded file using MD5	
DATA	`technology`	Enum	Required	Technology.: Technology used for the assay. e.g. 'sN' for single Nuclei or 'sC' for single Cell.	['SN','SC','Spatial','Bulk','Seq','Array']
DATA	`omic`	Enum	Required	Omic.: Readout type. This will be "RNA" for most of the ASAP CRN data.	['RNA','DNA','Hi-C','ATAC','Protein']
DATA	`adjustment`	Enum	Required	Adjustment. : Indication of whether the data has been adjusted. i.e. Is the data "raw" or "processed"?	['Raw','Processed']
DATA	`content`	Enum	Required	Content.: What is the variable measurement values?	['Counts','Probabilities','Genotypes','Dosages','Reads']
DATA	`time`	Float	Required	Time.: Time from baseline in days. Units in days, 0 is baseline to facilitate longitudinal analysis.	
DATA	`header`	String	Required	Header.: What are the primary header labels?	
DATA	`annotation`	String	Required	Annotation.: Additional annotation columns present in file. I.e. 'tissue','UMI','barcode','cluster', or'cell location' ... Additional colunns in the annData or HDF files when relevant to technologies like single cell or spatial. ['SNP','ENSG','interval'] SNP is rsID or positional annotation. Interval is in chr_start_stop format. ENSG is ensemble ID.	
DATA	`configuration_file`	String	Required	Configuration File.: Incl. software version, reference assembly, parameters for analysis, in YAML format	

CLINPATH

Table	Column Name	Data Type	Required	Description	Enum Values
CLINPATH	subject_id	String	Required	Subject ID.: Can be multiple samples from each Subject	
CLINPATH	source_subject_id	String	Required	Source Sample ID.: ID provided by Brain bank or other source	
CLINPATH	duration_pmi	Float	Required	Duration of Pmi.: Post Mortem Interval (PMI) between death and autopsy start. Units in hours	
CLINPATH	path_autopsy_dx_main	Enum	Required	Pathology Autopsy Diagnosis (Main): The primary pathological autopsy result. Code general healthy control as: "Control, no misfolded protein or significant vascular pathology"	["Lewy body disease nos", "Parkinson's disease", "Parkinson's disease with dementia", "Dementia with Lewy bodies", "Multiple system atrophy (SND>OPCA)", "Multiple system atrophy (OPCA<SND)", "Multiple system atrophy (SND=OPCA)", "Progressive supranuclear palsy", "Corticobasal degeneration", "Globular glial tauoapathy (GGT)", "Chronic traumatic encephalopathy (CTE)", "FTLD-Tau (Pick's)", "FTLD-Tau (MAPT)", "FTLD-Tau (AGD)", "FTLD-TDP43, Type A", "FTLD-TDP43, Type B", "FTLD-TDP43, Type C", "FTLD-TDP43, Type D", "FTLD-TDP43, Type E", "Motor neurone disease-TDP43 (MND or ALS)", "FTLD-MND-TDP43", "Huntington's disease", "Spinocerebellar ataxia, nos", "Prion disease, nos", "Alzheimer's disease (high level neuropathological change)", "Alzheimer's disease (intermediate level neuropathological change)", "Control, Low level AD neuropathological change", "Control,

					Limbic predominant age-related TDP43 proteinopathy (LATE)", "Control, Argyrophilic grain disease", "Control, Primary age-related tauopathy (PART)", "Control, Ageing-related tau astrogliopathy (ARTAG)", "Control, Cerebrovascular disease (atherosclerosis)", "Control, Cerebrovascular disease (hyaline arteriolosclerosis)", "Control, Cerebrovascular disease (cerebral amyloid angiopathy)", "Control, no misfolded protein or significant vascular pathology", "Other neurological disorder"]
CLINPATH	path_autopsy_second_dx	String	Optional	Pathology Autopsy Diagnosis (2): The second listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_autopsy_third_dx	String	Optional	Pathology Autopsy Diagnosis (3): The third listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_autopsy_fourth_dx	String	Optional	Pathology Autopsy Diagnosis (4): The fourth listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_autopsy_fifth_dx	String	Optional	Pathology Autopsy Diagnosis (5): The fifth listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	

CLINPATH	path_autopsy_sixth_dx	String	Optional	Pathology Autopsy Diagnosis (6): The sixth listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_autopsy_seventh_dx	String	Optional	Pathology Autopsy Diagnosis (7): The seventh listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_autopsy_eight_dx	String	Optional	Pathology Autopsy Diagnosis (8): The eighth listed neuropathological diagnosis as provided by the brain bank for the donor from which the sample is acquired; if none listed, state "none"	
CLINPATH	path_year_death	Float	Required	Year of Death.: Year in which the donor from which the sample is derived died. Must be between 1920 and 2050	(y>1920)&(y<2050)
CLINPATH	age_at_death	String	Required	Age at Death.: Age in years when the donor from which the sample is derived died	
CLINPATH	cause_death	String	Required	Cause of Death.: The direct cause of death.	
CLINPATH	other_cause_death_1	String	Optional	Other Cause of Death (1): Additional (secondary) causes of death; if not available, to state "unknown"	
CLINPATH	other_cause_death_2	String	Optional	Other Cause of Death (2): Additional (tertiary) causes of death; if not available, to state "unknown"	
CLINPATH	brain_weight	Float	Required	Brain Weight.: Post-mortem brain weight in grams.	
CLINPATH	path_braak_nft	Enum	Required	Pathology Braak NFT.: Braak&Braak neurofibrillary tangle tau stage.	["0", "I", "II", "III", "IV", "V", "VI", "I/II", "III/IV", "IV/V", "V/VI"]
CLINPATH	path_braak_asyn	Enum	Required	Pathology Braak asyn: Braak&Braak lewie body stage.	["0", "1", "2", "3", "4", "5", "6", "1/2", "3/4", "4/5", "5/6"]
CLINPATH	path_cerad	Enum	Required	Pathology CERAD: Semiquantitative assessment of neuritic plaques.	["None", "Sparse", "Moderate", "Frequent"]

CLINPATH	path_thal	Enum	Required	Pathology Thal : Thal stage of Amyloid beta.	["0", "1", "2", "3", "4", "5", "1/2", "3", "4/5"]
CLINPATH	known_pathogenic_mutation	Enum	Optional	Known Pathogenic Mutation.: Does the donor from which the sample is acquired is known to have a pathogenic Parkinson's disease -causing mutation as known to the brain bank, including GBA?	["None", "Present", "Unknown"]

CLINPATH	PD_pathogenic_mutation	String	Optional	PD Pathogenic Mutation.: To specify as a free text which pathogenic mutation the donor has (e.g., LRRK2 G2019S, GBA L444P)	
CLINPATH	path_mckeith	Enum	Optional	Pathology McKeith.: McKeith and USSLBD Stage.	["Neocortical", "Limbic (transitional)","Brainstem", "Amygdala Predominant", "Olfactory bulb only","Limbic, transitional (brainstem and limbic involvement)", "Diffuse, neocortical (brainstem, limbic and neocortical involvement)", "Olfactory Bulb-Only", "Limbic (amygdala) predominant", "Absent", "Present, but extent unknown"]
CLINPATH	sn_neuronal_loss	Enum	Optional	SN Neuronal Loss.: sn_neuronal_loss (none, mild, moderate, severe, 0-3)	["None", "Mild", "Moderate", "Severe", "Not assessed", "Unknown"]
CLINPATH	path_infarcs	Enum	Optional	Pathology Infarcs.: Cerebral infarcts (indicator for a large infarct or in study region)	["Yes", "No"]
CLINPATH	path_nia_ri	Enum	Optional	Pathology NIA_RI.: NIA_RI (Reagan Institute) criteria: 'low', 'intermediate', 'high', 'none'	['Low', 'Intermediate', 'High', 'None']
CLINPATH	path_nia_aa_a	Enum	Optional	Pathology NIA-AA A.: NIA-AA A criteria (pathological not biomarker criteria)	["A0","A1","A2","A3"]
CLINPATH	path_nia_aa_b	Enum	Optional	Pathology NIA-AA B.: NIA-AA B criteria (pathological not biomarker criteria)	["B0","B1","B2","B3"]
CLINPATH	path_nia_aa_c	Enum	Optional	Pathology NIA-AA C.: NIA-AA C criteria (pathological not biomarker criteria):	["C0","C1","C2","C3"]
CLINPATH	TDP43	Enum	Optional	TDP43.: Specifies if concomitant TDP43 pathology is present in the medial temporal and frontal lobes; this information to be provided by the brain bank, based on the examination of formalin fixed paraffin embedded tissue of the whole or half of the brain.	["None in medial temporal lobe", "Present in amygdala, only", "Present in hippocampus, only", "Present in amygdala and hippocampus, only", "Present in medial temporal lobe and middle frontal gyrus (not FTLD pattern)", "Unknown"]
CLINPATH	arteriolosclerosis_severity_scale	Enum	Optional	Arteriolosclerosis Severity Scale: Arteriolosclerosis severity scale.	["None", "Mild", "Moderate", "Severe", "Not assessed", "Unknown"]

CLINPATH	amyloid_angiopathy_severity_scale	Enum	Optional	Amyloid Angiopathy Severity Scale: Amyloid angiopathy severity scale.	["None", "Mild", "Moderate", "Severe", "Not assessed", "Unknown"]
CLINPATH	path_ad_level	Enum	Optional	Patholgy AD Level.: Specifies the level of Alzheimer's disease neuropathological change, based on 2012 NIA-AA criteria using the options in the dropdown menu: this information to be provided by the brain bank, based on the examination of the whole or half of the brain.	["No evidence of Alzheimer's disease neuropathological change", "Low level Alzheimer's disease neuropathological change", "At least low level Alzheimer's disease neuropathological change", "Intermediate level Alzheimer's disease neuropathological change", "At least intermediate level Alzheimer's disease neuropathological change", "High level Alzheimer's disease neuropathological change", "Unknown"]
CLINPATH	dig_slide_avail	Enum	Optional	Digital Slide Available: Indicates if a digital image of the pathology slides are available.	["Yes", "No"]
CLINPATH	quant_path_avail	Enum	Optional	Quantitative Pathology Available: Indicates if quantative pathology is available.	["Yes", "No"]