Change log overview: PDS1-PDS6

Contents

Overview

This document provides an overview of changes made to the structure, content, and format of the Enroll HD Periodic Dataset (PDS) between each consecutive major (PDSX) and minor (RX) release, from the first (PDS1; R1) to the most recent (PDS6; R2). Information about files and file structure, variables (additions, exclusions, and modifications), coding systems, and quality control and identification risk protocol is provided. Sample size and visit counts by major and minor release are also documented. Individual participant value changes between major and minor releases are captured in separate documents, available on request.

Files and file structure

The files contained in each Enroll-HD PDS release are displayed in Table 1.

Table 1 - Files included in each PDS release.

		PDS release						
		PDS1 *R1*	*PDS2* *R1, R2*	*PDS3*		*PDS4* *R1, R2, R3*	*PDS5* *R1, R2*	*PDS6* *R1, R2*
				R1	*R2*			
File	adhoc	-	-	✓	✓	✓	✓	✓
	assessment	-	-	✓	✓	✓	✓	✓
	baseline	✓	✓	-	-	-	-	-
	comorbid	✓	✓	✓	✓	✓	✓	✓
	enroll	-	-	✓	✓	✓	✓	✓
	event	-	-	✓	✓	✓	✓	✓
	followup	✓	✓	-	-	-	-	-
	general	✓	✓	-	-	-	-	-
	nonpharmacotx	✓	✓	✓	✓	✓	✓	✓
	nutsuppl	✓	✓	✓	✗	✓	✓	✓
	participation	-	-	✓	✓	✓	✓	✓
	pharmacotx	✓	✓	✓	✗	✓	✓	✓
	phonecontact	✓	✓	-	-	-	-	-
	prematureend	-	✓	-	-	-	-	-
	profile	-	-	✓	✓	✓	✓	✓
	registry	-	-	✓	✓	✓	✓	✓
	subject	✓	✓	-	-	-	-	-

PDS1 contained 9 files: *baseline*, *comorbid*, *followup*, *general*, *nonpharmacotx*, *nutsuppl*, *pharmacotx*, *phonecontact*, *and subject*. For PDS2, the *prematureend* file was added to the release.

The main changes in the PDS release file structure occurred between PDS2 and PDS3; seven new files were added (*adhoc*, *assessment*, *enroll*, *event*, *participation*, *profile*, *registry*), and six files were removed (*baseline*, *followup*, *general*, *phonecontact*, *prematureend*, *subject*).

PDS1 and PDS2 contained only data from the Enroll-HD study. PDS3, and all subsequent releases, also contain data from Registry 2, Registry 3 and Adhoc study data, where these data are available for Enroll-HD participants. As a result, the *registry* and *adhoc* files were added for PDS3 and onwards. The *registry* file contains visit information collected as part of the Registry 2 and Registry 3 studies. All the visit information is included in a single file and includes information collected at baseline, each follow-up visit, and unscheduled visits. The *adhoc* file contains data collected about Enroll-HD participants prior to their participation in the Registry and Enroll HD studies. Both *registry* and *adhoc* files share the same structure as the *enroll* file, which was also introduced in PDS3 to capture the data previously included in the *baseline*, *followup*, and *phonecontact* files (see Figure 1). The *enroll* file also includes data from unscheduled visits that were not made available previously. The new *assessment* file summarizes all available assessments/forms for visits from each study; this file can be a very useful tool to determine which data are available from each study. The *participation* file (see Figure 1) contains all the variables from *prematureend* and some variables from the *subject* file (for example *hdcat_l* and *subjstat*). The *profile* file (see Figure 1) includes all the variables from the previously used *general* file and the additional variables presented in the *subject* file, namely *region*, *caghigh* and *caglow*.

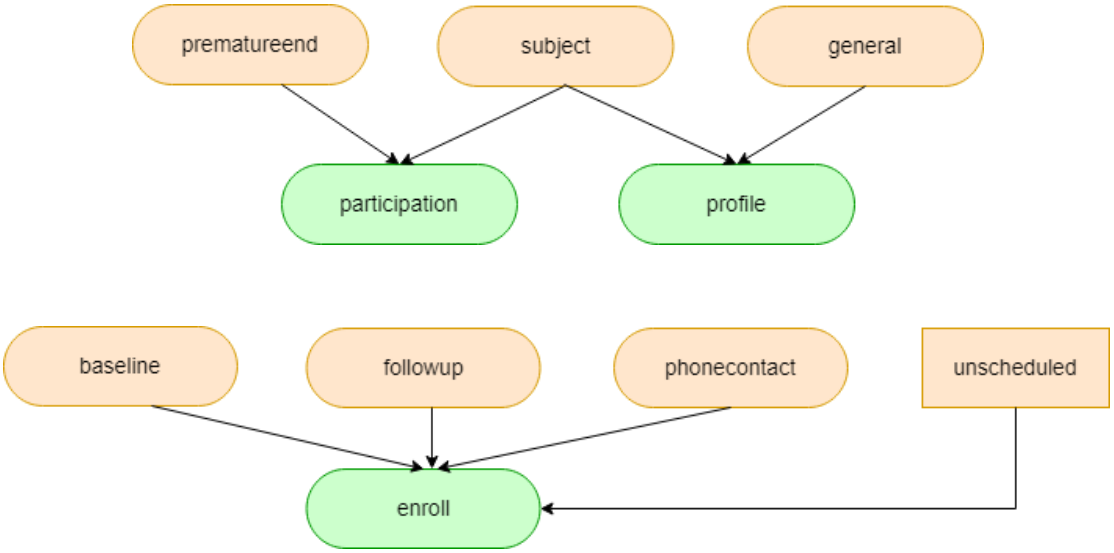

Figure 1 - Files undergoing structural changes only (same information retained) from PDS2 to PDS3. Orange boxes indicate PDS1 and PDS2 files; green boxes indicate files from PDS3 to the current PDS release.

Finally, the *event* file - an Enroll HD study specific file – was introduced for PDS3. This file was not available until PDS3 because the number of observed Enroll-HD reportable events prior to this release was relatively small, causing concerns about potential identification.

Four files remained the same in file name and structure since the beginning: *comorbid*, *pharmacotx*, *nonpharmacotx*, and *nutsuppl*. However, in the PDS3_R2 release, two files (*pharmacotx* and *nutsuppl*) were removed to limit the distribution of the MedDRA and WHO-DD coded data to researchers holding a MedDRA and/or WHO-DD license. For more details, please consult Variables and Coding sections.

No further modifications to file structure have been made since PDS3.

Variables

This section summarizes changes made to variables within each file between consecutive PDS releases, including new variables, variable exclusions, and variable modifications.

The number of variables included in each file of each PDS release is summarized in Table 2.

Table 2 - Number of variables by file and PDS release.

		PDS release					
		PDS1 *R1*	*PDS2* *R1, R2*	*PDS3* *R1*	*PDS3* *R2*	*PDS4* *R1, R2, R3*	*PDS5* *R1, R2*	*PDS6* *R1, R2*
File	adhoc	-	-	109		109	109	108
	assessment	-	-	33		34	34	34
	baseline	223	228	-		-	-	-
	comorbid	9	8	7		7	7	7
	enroll	-	-	296		304	305	308
	event	-	-	12		12	12	12
	followup	191	208	-		-	-	-
	general	45	45	-		-	-	-
	nonpharmacotx	9	8	7		7	7	6
	nutsuppl	13	9	12	-	12	12	12
	participation	-	-	56		55	55	55
	pharmacotx	17	15	14	-	14	14	14
	phonecontact	10	10	-		-	-	-
	prematureend	-	6	-		-	-	-
	profile	-	-	42		62	62	62
	registry	-	-	287		287	288	286
	subject	9	9	-		-	-	-

Tables Table *3*, Table *5*, Table *7*, Table *8*, and Table *9* provide information on sequential changes to variables (inclusions, exclusions, and modifications) between the final major versions of each PDS release.

Label	Description	Coding
New	Indicates a variable which was not included in the prior dataset release	
Removed	Indicates a variable which existed in prior dataset release(s) which has now been removed in totality	
Modified	Indicates a change to an existing variable (e.g., change in variable name, response options)	
Key variable	Indicates a variable necessary to combine data files (e.g., subjid) Key variable(s) must be included each (new) file.	❋

For variables changes implemented between minor releases (PDS2_R1 to PDS2_R2, PDS3_R1 to PDS3_R2, and PDS6_R1 to PDS6_R2), a brief description of the reason why those changes were made and the variables impacted are available. Table *4* and Table *6* resume the changes between the different minor releases of PDS2 and PDS3, respectively.

The majority of variable changes were observed between PDS2 and PDS3 when data from additional studies were added, and additional data from the Enroll-HD study were made available in the PDS (e.g., events).

PDS1_R1 → PDS2_R2

Table 3 - Changes in the variables between PDS1 and PDS2.

File	Variable	Label	Status
Baseline	occupant	UHDRS Total Functional Capacity - Occupation	
Baseline	finances	UHDRS Total Functional Capacity - Finances	
Baseline	chores	UHDRS Total Functional Capacity - Domestic chores	
Baseline	adl	UHDRS Total Functional Capacity - ADL	
Baseline	carelevl	UHDRS Total Functional Capacity - Care level	
Comorbid	seq	Visit Index	❋
Followup	seq	Visit Index	❋
Followup	occupant	UHDRS Total Functional Capacity - Occupation	
Followup	finances	UHDRS Total Functional Capacity - Finances	
Followup	chores	UHDRS Total Functional Capacity - Domestic chores	
Followup	adl	UHDRS Total Functional Capacity - ADL	
Followup	carelevl	UHDRS Total Functional Capacity - Care level	
Followup	pbas1wo	Group Depressed mood - Worst	
Followup	pbas2wo	Group Suicidal ideation - Worst	
Followup	pbas3wo	Group Anxiety - Worst	
Followup	pbas4wo	Group Irritability - Worst	

Followup	pbas5wo	Group Angry or aggressive behavior - Worst	
Followup	pbas6wo	Group Lack of initiative (apathy) - Worst	
Followup	pbas7wo	Group Perseverative thinking or behavior - Worst	
Followup	pbas8wo	Group Obsessive-Compulsive Behaviours - Worst	
Followup	pbas9wo	Group Delusions / paranoid thinking - Worst	
Followup	pbas10wo	Group Hallucinations - Worst	
Followup	pbas11wo	Group Disoriented Behaviour - Worst	
Nonpharmacotx	seq	Visit Index	❋
Nutsuppl	cmdostotl	Total daily dose	
Nutsuppl	cmdosfrq	Frequency	
Nutsuppl	seq	Visit Index	❋
Nutsuppl	cmdose__cmdostxt	Dose / Unit (dose of one intake)	
Nutsuppl	cmfrq__1	Daily intake - in the morning	
Nutsuppl	cmfrq__2	Daily intake - at noon	
Nutsuppl	cmfrq__3	Daily intake - in the evening	
Nutsuppl	cmfrq__4	Daily intake - at night	
Pharmacotx	cmtrt__ing	Ingredient – Modified Term	
Pharmacotx	cmtrt__atc	Ingredient – Code (coded by ATC)	
Pharmacotx	cmdostot	Total daily dose	
Pharmacotx	cmdosfrq	Frequency	
Pharmacotx	seq	Visit Index	❋
Pharmacotx	cmdose__cmdostx	Dose	
Pharmacotx	cmfrq__1	Daily intake - in the morning	
Pharmacotx	cmfrq__2	Daily intake - at noon	
Pharmacotx	cmfrq__3	Daily intake - in the evening	
Pharmacotx	cmfrq__4	Daily intake - at night	
Phonecontact	seq	Visit Index	❋
Phonecontact	dsterm	Specify primary reason for participant's premature discontinuation from study (lost to follow up hint: failure of participant to return to follow-up visit and failure.	
Prematureend	subjid	HDID (recoded)	❋
Prematureend	studyid	Study ID	❋
Prematureend	visit	Visit name	❋
Prematureend	visdy	Visit day	❋
Prematureend	visstat	Visit status	❋
Prematureend	dsterm	Specify primary reason for participant's Premature discontinuation from study (lost to follow up hint: failure of participant to return to follow-up visit and failure to be located by investigator)	
Subject	region	Region (latest)	❋

Variable notes, including modification(s):

- The option 'Oceania' was changed to 'Australasia' in the variable *region*.

.

Variable changes for PDS2 releases (R1, R2):

Regarding the releases (R1, R2) of the PDS2, just one variable change was made: the variable *mhenr* was renamed as *mhenrf* on the *comorbid* file. The variable coding remained unchanged, only the name was corrected and matches the name released on PDS1. The variables changes between releases are presented in Table 4.

Table 4 - Variables with changes between PDS2 minor releases

File	Variable		Label	Reason
	PDS2_R1	*PDS2_R1*		
comorbid	mhenr	*mhenrf*	Ongoing	Name corrected

PDS2_R2 → PDS3_R2

Table 5 - Changes in the variables between PDS2 and PDS3.

File	Variable	Label	Status
Adhoc	age	Age at visit	❋
Adhoc	emplusl	Could subject engage in gainful employment in his/her accustomed work?	
Adhoc	emplany	Could subject engage in any kind of gainful employment?	
Adhoc	volunt	Could subject engage in any kind of volunteer or nongainful work?	
Adhoc	fafinan	Could subject manage his/her finances (monthly) without any help?	
Adhoc	grocery	Could subject shop for groceries without help?	
Adhoc	cash	Could subject handle money as a purchaser in a simple cash (shop) transaction?	
Adhoc	supchild	Could subject supervise children without help?	
Adhoc	drive	Could subject operate an automobile safely and independently?	

Adhoc	housework	Could subject do his/her own housework without help?	
Adhoc	laundry	Could subject do his/her own laundry (wash/dry) without help?	
Adhoc	prepmeal	Could participant prepare his/her own meals without help?	
Adhoc	telephon	Could subject use the telephone without help?	
Adhoc	ownmeds	Could subject take his/her own medications without help?	
Adhoc	feedself	Could subject feed himself/herself without help?	
Adhoc	dress	Could subject dress himself/herself without help?	
Adhoc	bathe	Could subject bathe himself/herself without help?	
Adhoc	pubtrans	Could subject use public transport to get to places without help?	
Adhoc	walknbr	Could subject walk to places in his/her neighbourhood without help?	
Adhoc	walkfall	Could subject walk without falling?	
Adhoc	walkhelp	Could subject walk without help?	
Adhoc	comb	Could subject comb hair without help?	
Adhoc	trnchair	Could subject transfer between chairs without help?	
Adhoc	bed	Could subject get in and out of bed without help?	
Adhoc	toilet	Could subject use toilet/commode without help?	
Adhoc	carehome	Could subject's care still be provided at home?	
Adhoc	hvlt	Hopkins Verbal Learning Test-R completed (optional)	
Assessment	subjid	HDID (recoded)	❋
Assessment	studyid	Study ID	❋
Assessment	visit	Visit name	❋
Assessment	seq	Visit index	❋
Assessment	visdy	Visit day	❋
Assessment	enrollment	Enrollment	
Assessment	variable	Variable	
Assessment	mhx	Medical History	
Assessment	samples	Samples	
Assessment	motor	UHDRS Motor/Diagnostic Confidence	
Assessment	tfc	UHDRS Total Functional Capacity	
Assessment	funct	UHDRS Function Assessment/Independence Scale	
Assessment	cognitive	Cognitive Assessments	
Assessment	mmse	Mini Mental State Examination	
Assessment	physio	Physiotherapy Outcome Measures	
Assessment	pbas	Problem Behaviours Assessment - Short	
Assessment	csri	Client Service Receipt Inventory	
Assessment	careqol	Caregivers Quality of Life Questionnaire	

Assessment	sf12	Short Form Health Survey – 12v2	
Assessment	hadssis	Hospital Anxiety and Depression Scale – Snaith Irritability Scale	
Assessment	cssrsfup	Columbia-Suicide Severity Rating Scale - Annual	
Assessment	cssrsbl	Columbia-Suicide Severity Rating Scale - Baseline	
Assessment	wpaishp	WPAI-SHP	
Assessment	missedvisit	Missed Visit	
Assessment	behaviour	UHDRS Behavioral Assessment	
Assessment	summary	Summary	
Assessment	gci	Global Clinical Impression Scale	
Assessment	care	Care Giver Questionnaire	
Assessment	pbahd	Problem Behaviours Assessment for HD	
Assessment	sf36v2	SF-36v2 Health Survey	
Assessment	bdi	Becks Depression Scale	
Assessment	hamilton	Hamilton Rating Scale for Depression	
Assessment	sf36v1	SF-36v1 Health Survey	
Comorbid	studyid	Study ID	*
Enroll	age	Age at visit	*
Enroll	emplusl	Could subject engage in gainful employment in his/her accustomed work?	
Enroll	emplany	Could subject engage in any kind of gainful employment?	
Enroll	volunt	Could subject engage in any kind of volunteer or nongainful work?	
Enroll	fafinan	Could subject manage his/her finances (monthly) without any help?	
Enroll	grocery	Could subject shop for groceries without help?	
Enroll	cash	Could subject handle money as a purchaser in a simple cash (shop) transaction?	
Enroll	supchild	Could subject supervise children without help?	
Enroll	drive	Could subject operate an automobile safely and independently?	
Enroll	housework	Could subject do his/her own housework without help?	
Enroll	laundry	Could subject do his/her own laundry (wash/dry) without help?	
Enroll	prepmeal	Could participant prepare his/her own meals without help?	
Enroll	telephon	Could subject use the telephone without help?	
Enroll	ownmeds	Could subject take his/her own medications without help?	
Enroll	feedself	Could subject feed himself/herself without help?	
Enroll	dress	Could subject dress himself/herself without help?	

Enroll	bathe	Could subject bathe himself/herself without help?	
Enroll	pubtrans	Could subject use public transport to get to places without help?	
Enroll	walknbr	Could subject walk to places in his/her neighbourhood without help?	
Enroll	walkfall	Could subject walk without falling?	
Enroll	walkhelp	Could subject walk without help?	
Enroll	comb	Could subject comb hair without help?	
Enroll	trnchair	Could subject transfer between chairs without help?	
Enroll	bed	Could subject get in and out of bed without help?	
Enroll	toilet	Could subject use toilet/commode without help?	
Enroll	carehome	Could subject's care still be provided at home?	
Enroll	sid1	Have you wished you were dead or wished you could go to sleep and not wake up?	
Enroll	sid2	Have you actually had any thoughts of killing yourself	
Enroll	sid3	Have you been thinking about how you might do this?	
Enroll	sid4	Have you had these thoughts and had some intention of acting on them?	
Enroll	sid5	Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?	
Enroll	int1	Group Intensity of Ideation – Most Severe – Type	
Enroll	int2	How many times have you had these thoughts?	
Enroll	int3	When you have the thoughts, how long do they last?	
Enroll	int4	Could/can you stop thinking about killing yourself or wanting to die if you want do?	
Enroll	int5	Are there things – anyone or anything (e.g. family, religion, pain of death) – that stopped you from wanting to die or acting on thoughts of committing suicide?	
Enroll	int6	What sort of reasons did you have for thinking about wanting to die or killing yourself?	
Enroll	sbh1	Group Suicidal Behavior - Actual attempt	
Enroll	sbh1n	Total # of attempts	
Enroll	sbh2	Has subject engaged in Non-Suicidal Self-Injurious Behavior?	
Enroll	sbh3	Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?	
Enroll	sbh3n	Total # of interrupted	

Enroll	sbh4	Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?	
Enroll	sbh4n	Total # of aborted	
Enroll	sbh5	Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?	
Enroll	sbh6	Suicidal behaviour was present during the assessment period?	
Enroll	sbh7	Completed Suicide was present during the assessment period	
Enroll	attmpt1dy	Day of most recent attempt	
Enroll	attmpt11	Actual Lethality/Medical Damage	
Enroll	attmpt12	Potential Lethality	
Enroll	attmpt2dy	Day of most lethal attempt	
Enroll	attmpt21	Actual Lethality/Medical Damage	
Enroll	attmpt22	Potential Lethality	
Enroll	attmpt3dy	Day of Initial/First attempt	
Enroll	attmpt31	Actual Lethality/Medical Damage	
Enroll	attmpt32	Potential Lethality	
Enroll	rfstdy	Enrollment day	
Enroll	fhx	Family History	
Enroll	bsp	Biosamples for use in research	
Enroll	nvl	Participation in sub-studies	
Enroll	cntct	Contact between visits	
Enroll	cntctst	Contact regarding other research studies	
Enroll	cntctpmt	Contact regarding post-mortem tissue collection	
Enroll	rtrspctv	Linking clinical information from previous studies	
Event	subjid	HDID (recoded)	❋
Event	studyid	Study ID	❋
Event	seq	Index	❋
Event	evtdy	Day of report	
Event	evtcode	Reportable event codes	
Event	evtterm__modify	Event term – Modified Term (coded with MedDRA)	
Event	evtterm__decod	Event term – Code (coded with MedDRA)	
Event	evtstdy	Onset day	
Event	stdtcest	Date estimated	
Event	evtongo	Resolved	
Event	evtendy	End Day	
Event	evtdsmc	Status of DSMC review	
Nonpharmacotx	studyid	Study ID	❋

Nutsuppl	subjid	HDID (recoded)	*
Nutsuppl	studyid	Study ID	*
Nutsuppl	cmcat	Type	
Nutsuppl	cmdostot	Total daily dose	
Nutsuppl	cmdosunit	Unit	
Nutsuppl	cmdosfrq	Frequency	
Nutsuppl	cmstdy	Start Day	
Nutsuppl	cmenrf	Ongoing	
Nutsuppl	cmendy	End day	
Participation	study	Study	*
Participation	subjstat	Participant status (latest)	
Participation	hdcat_0	Participant category (at enrollment)	
Participation	hdcat_l	Participant category (latest)	
Participation	age_0	Age at enrollment	
Participation	rfstdy	Study reference start day (baseline)	
Participation	rficdy	Day of informed consent	
Participation	rfendy	Day of end of study	
Participation	evtnum	Number of events	
Participation	saenum	Number of serious events	
Participation	visitnum	Number of visits	
Participation	visit1, …, visit21	Visit type	
Participation	vis1dy,…, vis21dy	Day of visit	
Participation	seq	Visit Index	*
Participation	visit	Visit name	*
Participation	visdy	Visit day	*
Participation	visstat	Visit Status	*
Pharmacotx	subjid	HDID (recoded)	*
Pharmacotx	studyid	Study ID	*
Pharmacotx	cmtrt__modify	Drug name – Modified Term (coded by WHO-DD)	
Pharmacotx	cmtrt__decod	Drug name – Code (coded by WHO-DD)	
Pharmacotx	cmtrt__ing	Drug name - Ingredients	
Pharmacotx	cmtrt__atc	Drug name – ATC Codes	
Pharmacotx	cmindc__modify	Indication – Modified Term (coded by MedDRA)	
Pharmacotx	cmindc__decod	Indication – Code (coded by MedDRA)	
Pharmacotx	cmdostot	Total daily dose	
Pharmacotx	cmdose__cmdosu	Unit (unit of one dose)	
Pharmacotx	cmdosfrq	Frequency	
Pharmacotx	cmroute	Route	
Pharmacotx	cmstdy	Start day	
Pharmacotx	cmenrf	Ongoing	
Pharmacotx	cmendy	End day	
Profile	studyid	Study ID	*

Profile	visit	Visit name	❋
Profile	visdy	Visit day	❋
Profile	visstat	Visit status	❋
Profile	age	Age at visit	❋
Profile	hdtest	Has an HD genetic test been done?	
Profile	lbdy	Collection Day	
Profile	fhx	Family History	
Registry	age	Age at visit	❋
Registry	emplusl	Could subject engage in gainful employment in his/her accustomed work?	
Registry	emplany	Could subject engage in any kind of gainful employment?	
Registry	volunt	Could subject engage in any kind of volunteer or nongainful work?	
Registry	fafinan	Could subject manage his/her finances (monthly) without any help?	
Registry	grocery	Could subject shop for groceries without help?	
Registry	cash	Could subject handle money as a purchaser in a simple cash (shop) transaction?	
Registry	supchild	Could subject supervise children without help?	
Registry	drive	Could subject operate an automobile safely and independently?	
Registry	housework	Could subject do his/her own housework without help?	
Registry	laundry	Could subject do his/her own laundry (wash/dry) without help?	
Registry	prepmeal	Could participant prepare his/her own meals without help?	
Registry	telephon	Could subject use the telephone without help?	
Registry	ownmeds	Could subject take his/her own medications without help?	
Registry	feedself	Could subject feed himself/herself without help?	
Registry	dress	Could subject dress himself/herself without help?	
Registry	bathe	Could subject bathe himself/herself without help?	
Registry	pubtrans	Could subject use public transport to get to places without help?	
Registry	walknbr	Could subject walk to places in his/her neighbourhood without help?	
Registry	walkfall	Could subject walk without falling?	
Registry	walkhelp	Could subject walk without help?	
Registry	comb	Could subject comb hair without help?	
Registry	trnchair	Could subject transfer between chairs without help?	
Registry	bed	Could subject get in and out of bed without help?	

Registry	toilet	Could subject use toilet/commode without help?	
Registry	carehome	Could subject's care still be provided at home?	
Registry	hvlt	Hopkins Verbal Learning Test-R completed (optional)	
Registry	hvltd	Used form	
Registry	hvltt11	Trial 1: correct	
Registry	hvltt21	Trial 2: correct	
Registry	hvltt31	Trial 3: correct	
Registry	hvlt1	Delayed recall correct	
Registry	hvlt2	Total repetitions (trials 1-3 & delayed)	
Registry	hvlt3	Total intrusions (trials 1-3 & delayed)	
Registry	hvlt4	Recognition: true positives	
Registry	hvlt5	Recognition: false positives	
Registry	mdrs	Dementia Rating Scale-2 completed (optional)	
Registry	mdrs1	Attention	
Registry	mdrs2	Initiation / perseveration	
Registry	mdrs3	Construction	
Registry	mdrs4	Conceptualization	
Registry	mdrs5	Memory	
Registry	sfscore_v1	SF-36 score	
Registry	sfscore_v2	SF-36 v.2 score	
Registry	behscore	Behavioral score	
Registry	biscore	Behavioral score incomplete	
Registry	behaviour_depscore	Group Sub-scores - Depression	
Registry	defscore	Group Sub-scores - Drive/executive function	
Registry	behaviour_irascore	Group Sub-scores - Irritability/aggression	
Registry	behaviour_psyscore	Group Sub-scores - Psychosis	
Registry	deprfr	Group Depressed mood - Frequency	
Registry	deprsv	Group Depressed mood - Severity	
Registry	guiltfr	Group Low self-esteem/guilt - Frequency	
Registry	guiltsv	Group Low self-esteem/guilt - Severity	
Registry	anxifr	Group Anxiety - Frequency	
Registry	anxisv	Group Anxiety - Severity	
Registry	suicfr	Group Suicidal thoughts - Frequency	
Registry	suicsv	Group Suicidal thoughts - Severity	
Registry	aggrfr	Group Disruptive or aggressive behavior - Frequency	
Registry	aggrsv	Group Disruptive or aggressive behavior - Severity	
Registry	irritfr	Group Irritable behavior - Frequency	
Registry	irritsv	Group Irritable behavior - Severity	

Registry	obsessfr	Group Perseverative/obsessional thinking - Frequency	
Registry	obsesssv	Group Perseverative/obsessional thinking - Severity	
Registry	compulfr	Group Compulsive behavior - Frequency	
Registry	compulsv	Group Compulsive behavior - Severity	
Registry	delusfr	Group Delusions - Frequency	
Registry	delussv	Group Delusions - Severity	
Registry	hallucfr	Group Hallucinations - Frequency	
Registry	hallucsv	Group Hallucinations - Severity	
Registry	apathfr	Group Apathy - Frequency	
Registry	apathsv	Group Apathy - Severity	
Registry	confus	Does the examiner believe the participant is confused?	
Registry	dement	Does the examiner believe the participant is demented?	
Registry	deprexam	Does the examiner believe the participant is depressed?	
Registry	deprrfr	Does the participant require pharmacotherapy for depression?	
Registry	irritrfr	Does the participant require pharmacotherapy for irritability?	
Registry	isobt	Was the behavioral assessment information obtained from:	
Registry	bdiscore	Becks Score	
Registry	hamscore	Hamilton Score	
Registry	sid1	Have you wished you were dead or wished you could go to sleep and not wake up?	
Registry	sid2	Have you actually had any thoughts of killing yourself	
Registry	sid3	Have you been thinking about how you might do this?	
Registry	sid4	Have you had these thoughts and had some intention of acting on them?	
Registry	sid5	Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?	
Registry	int1	Group Intensity of Ideation – Most Severe – Type	
Registry	int2	How many times have you had these thoughts?	
Registry	int3	When you have the thoughts, how long do they last?	
Registry	int4	Could/can you stop thinking about killing yourself or wanting to die if you want do?	

Registry	int5	Are there things – anyone or anything (e.g. family, religion, pain of death) – that stopped you from wanting to die or acting on thoughts of committing suicide?	
Registry	int6	What sort of reasons did you have for thinking about wanting to die or killing yourself?	
Registry	sbh1	Group Suicidal Behavior - Actual attempt	
Registry	sbh1n	Total # of attempts	
Registry	sbh2	Has subject engaged in Non-Suicidal Self-Injurious Behavior?	
Registry	sbh3	Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?	
Registry	sbh3n	Total # of interrupted	
Registry	sbh4	Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?	
Registry	sbh4n	Total # of aborted	
Registry	sbh5	Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?	
Registry	sbh6	Suicidal behaviour was present during the assessment period?	
Registry	sbh7	Completed Suicide was present during the assessment period	
Registry	attmpt1dy	Day of most recent attempt	
Registry	attmpt11	Actual Lethality/Medical Damage	
Registry	attmpt12	Potential Lethality	
Registry	attmpt2dy	Day of most lethal attempt	
Registry	attmpt21	Actual Lethality/Medical Damage	
Registry	attmpt22	Potential Lethality	
Registry	attmpt3dy	Day of Initial/First attempt	
Registry	attmpt31	Actual Lethality/Medical Damage	
Registry	attmpt32	Potential Lethality	
Registry	svrtyll	Based upon my review of the information, I feel the severity of illness is	

Variable notes, including modification(s):

- All the variables of the *assessment* file were introduced in the PDS3 release, since this file was not available before;

- All the variables of the *event* file were introduced in the PDS3 release, since this file was not released on PDS2;
- The following assessments were introduced due to the inclusion of the Registry 2, Registry 3 and Retro-Visit: Short Form Health Survey – 36v2 (SF-36v2), Short Form Health Survey – 36v1 (SF-36v1), Beck – Depression Inventory (BDI), Behaviour, Hamilton, GCI, and C-SSRS (Baseline and Follow Up). Some of these assessments were also introduced into Enroll;
- The remaining variables of the *adhoc* and *registry* files were previously available in one of the visit files (*baseline*, *followup* or *phonecontact*) and, for that reason, are not included in the table.

Variable changes for PDS3 minor releases (R1 and R2):

Regarding the minor releases (R1, R2) of PDS3, two files previously made available in PDS3_R1 were removed in PDS3_R2, *nutsuppl* and *pharmacotx*, to limit distribution to only researchers holding a license for MedDRA and/or WHO-DD. The variables included in the *nutsuppl* and *pharmacotx* files on the PDS3_R1 and removed on PDS3_R2 are listed on Table 6.

Table 6 - Variables removed between PDS3 minor releases.

File	Variable	Label
Nutsuppl	subjid	HDID (recoded)
Nutsuppl	cmcat	Type
Nutsuppl	cmtrt__modify	Drug name – Modified Term (coded by WHO-DD)
Nutsuppl	cmtrt__decod	Drug name – Code (coded by WHO-DD)
Nutsuppl	cmtrt__atc	Drug name – ATC Codes
Nutsuppl	cmtrt__ing	Drug name – Ingredients
Nutsuppl	cmdostot	Total daily dose
Nutsuppl	cmdosunit	Unit
Nutsuppl	cmdosfrq	Frequency
Nutsuppl	cmstdy	Start day
Nutsuppl	cmenrf	Ongoing
Nutsuppl	cmendy	End day
Pharmacotx	subjid	HDID (recoded)
Pharmacotx	cmtrt__modify	Drug name – Modified Term (coded by WHO-DD)
Pharmacotx	cmtrt__decod	Drug name – Code (coded by WHO-DD)
Pharmacotx	cmtrt__ing	Drug name – Ingredients
Pharmacotx	cmtrt__atc	Drug name – ATC Codes

Pharmacotx	cmindc__modify	Indication – Modified Term (coded by MedDRA)
Pharmacotx	cmindc__decod	Indication – Code (coded by MedDRA)
Pharmacotx	cmdostot	Total daily dose
Pharmacotx	cmdose__cmdosu	Unit (unit of one dose)
Pharmacotx	cmdosfrq	Frequency
Pharmacotx	cmroute	Route
Pharmacotx	cmstdy	Start day
Pharmacotx	cmenrf	Ongoing
Pharmacotx	cmendy	End day

PDS3_R2 → PDS4_R3

Table 7 - Changes in the variables between PDS3 and PDS4.

File	Variable	Label	Status
Assessment	cognitivel	Extended Neuropsychological Assessment	
Assessment	function	UHDRS Function Assessment/Independence Scale	
Comorbid	mhterm__modify	Condition – Modified Term (coded)	
Comorbid	mhterm__decod	Condition – Code (coded)	
Enroll	hxtobcpd	Cigarettes per day	
Enroll	hxtobyos	Years of smoking	
Enroll	tobcpd	Cigarettes per day	
Enroll	tobyos	Years of smoking	
Enroll	sdmtnd	Symbol Digit Modality Test - Reason	
Enroll	verfctnd	Verbal Fluency Test – Reason	
Enroll	scntnd	Stroop Colour Naming Test - Reason	
Enroll	swrtnd	Stroop Word Reading Test - Reason	
Nutsuppl	subjid	HDID (recoded)	❋
Nutsuppl	cmcat	Type	
Nutsuppl	cmtrt__modify	NutSuppl Supplement – Modified Term	
Nutsuppl	cmtrt__decod	NutSuppl Supplement – Code	
Nutsuppl	cmtrt__atc	NutSuppl Supplement - ATC Code(s)	
Nutsuppl	cmtrt__ing	NutSuppl Supplement - ingredient(s)	
Nutsuppl	cmdostot	Total daily dose	
Nutsuppl	cmdosunit	Unit	
Nutsuppl	cmdosfrq	Frequency	
Nutsuppl	cmstdy	Start day	
Nutsuppl	cmenrf	Ongoing	
Nutsuppl	cmendy	End day	
Pharmacotx	subjid	HDID (recoded)	❋

Pharmacotx	cmtrt__modify	Drug name – Modified Term	
Pharmacotx	cmtrt__decod	Drug name - Code	
Pharmacotx	cmtrt__ing	Ingredient – Modified Term	
Pharmacotx	cmtrt__atc	Ingredient – Code (coded by ATC)	
Pharmacotx	cmindc__modify	Indication – Modified Term	
Pharmacotx	cmindc__decod	Indication - Code	
Pharmacotx	cmdostot	Total daily dose	
Pharmacotx	cmdose__cmdosu	Unit (unit of one intake)	
Pharmacotx	cmdosfrq	Frequency	
Pharmacotx	cmroute	Route	
Pharmacotx	cmstdy	Start day	
Pharmacotx	cmenrf	Ongoing	
Pharmacotx	cmendy	End day	
Participation	saenum	Number of serious events	
Pharmacotx	cmtrt__modify	Drug name – Modified Term	
Pharmacotx	cmtrt__decod	Drug name – Code	
Pharmacotx	cmindc__modify	Indication – Modified Term	
Pharmacotx	cmindc__decod	Indication – Code	
Profile	sxs_m	Group Mixed symptoms – motor	
Profile	sxs_c	Group Mixed symptoms – cognitive	
Profile	sxs_p	Group Mixed symptoms – psychiatric	
Profile	sxs_o	Group Mixed symptoms – oculomotor	
Profile	sxf_m	Group Mixed symptoms – motor	
Profile	sxf_c	Group Mixed symptoms – cognitive	
Profile	sxf_p	Group Mixed symptoms – psychiatric	
Profile	sxf_o	Group Mixed symptoms – oculomotor	
Profile	sxreas	Please specify why you, as a ratter, cannot estimate symptom onset (without additional external information) at the moment	
Profile	sxgs	What is your best guess of how many years ago symptom onset took place	
Profile	sxgsdy	Day of data entry	
Profile	sxr_m	Group Mixed symptoms – motor	
Profile	sxr_c	Group Mixed symptoms – cognitive	
Profile	sxr_p	Group Mixed symptoms – psychiatric	
Profile	sxr_o	Group Mixed symptoms – oculomotor	
Profile	xgwas	Additional GWAS data	
Profile	xbsp	Additional biosamples available	
Profile	xpheno	Additional phenotypic data available	
Profile	xmorpho	Additional morphometric data available	
Profile	ximage	Raw images available	

Variable notes and modification(s):

- The variable *funct* ("UHDRS Function Assessment/Independence Scale") of the *assessment* file had its name changed to *function* in PDS4 release;
- Regarding the *profile* file, additional variables (*xgwas*, *xbsp*, *xpheno*, *xmorpho*, and *ximage*) were added to allow researchers to easily identify if additional data or biosamples were available for each participant. These additional data and biosamples are from studies other than Enroll-HD and Registry, like Predict-HD, Track, among others, and can be requested through a SPS;
- Concerning the *pharmacotx* and *nutsupp* files that were reintroduced in the PDS4 release, all medications and nutritional supplements have been included using new Enroll-HD specific medication and comorbidity codes, which replace all MedDRA coding;
- Regarding the *comorbid* file, all WHO coding used in previous versions of PDS releases have been replaced by Enroll-HD specific codes.

Variable changes for PDS4 releases (R1, R2, R3):

Between the minor releases (R1, R2, and R3) of PDS4, there were no changes in the variables (included, excluded or modified). However, some coding on the following files were updated between PDS4_R1 and PDS4_R2: *comorbid*, *event*, *nutsuppl*, *participation*, *pharmacotx*, and *profile*:

- *profile*: correction of information to the variables *xgwas*, *xmorpho*, *xpheno*, *ximage*, and *xbsp*;
- *participation*: correction of the variable *rfendy*;
- *comorbid:* correction of *mhterm__modify* and *mhterm__decod;*
- *pharmacotx: correction of cmindc__decod;*
- *nutsuppl: correction of cmtrt__modify* and *cmtrt__decod*;
- *event*: correction of *evtterm__modify* and *evtterm__decod* .

PDS4_R3 → PDS5_R2

Table 8 - Changes in the variables between PDS4 and PDS5.

File	Variable	Label	Status
Enroll	updfh	Any updates to family history?	
Registry	updfh	Any updates to family history?	

Variable changes for PDS5 releases (R1, R2):

Between the minor releases (R1, R2) of the PDS5, there were no changes in the variables (included, excluded or modified).

PDS5_R2 → PDS6_R1

Table 9 - Changes in the variables between PDS5 and PDS6.

File	Variable	Label	Status
Adhoc	height	Height (cm)	
Adhoc	weight	Weight (kg)	
Adhoc	bmi	BMI	
Adhoc	bmi_imp	BMI imputed	
Adhoc	capscore	CAP score	
Enroll	hdiss_stage_imp	Imputed HD-ISS stage	
Enroll	hdiss_stage0_prob	Probability of classification as HD-ISS Stage 0	
Enroll	hdiss_stage1_prob	Probability of classification as HD-ISS Stage 1	
Enroll	hdiss_stage2_prob	Probability of classification as HD-ISS Stage 2	
Enroll	hdiss_stage3_prob	Probability of classification as HD-ISS Stage 3	
Enroll	capscore	CAP score	
Enroll	maristat	Marital status	
Enroll	height	Height (cm)	
Enroll	weight	Weight (kg)	
Enroll	bmi	BMI	
Enroll	bmi_imp	BMI imputed	
Nonpharmacotx	cmfrq	Daily intake	
Registry	maristat	Marital status	
Registry	height	Height (cm)	
Registry	weight	Weight (kg)	
Registry	bmi	BMI	
Registry	bmi_imp	BMI imputed	
Registry	capscore	CAP score	

Variable notes, including modification(s):

- The variable *bmi_imp* (BMI imputed) replaces the previous variable *bmi* and it is calculated based on observed weight at visit and height as observed at Enroll-HD baseline for all visits where the participant is at least 18 years of age (value is not provided for visits at which participant is < 18 years of age);
- The variable *capscore* (CAP score) was introduced – please refer to the **Understand and Interpret the Data** document for further information (see also: Warner JH, Long JD, Mills JA, Langbehn DR, Ware J, Mohan A, Sampaio C. Standardizing the CAP Score in Huntington's Disease by Predicting Age-at-Onset. J Huntingtons Dis. 2022;11(2):153-171. doi: 10.3233/JHD-210475. PMID: 35466943).
- Multiple variables were introduced concerning HD-ISS staging (*hdiss_stage_imp*, *hdiss_stage0_prob*, *hdiss_stage1_prob*, *hdiss_stage2_prob*, *hdiss_stage3_prob*) - please refer to the **Understand and Interpret the Data** document for further information (see also: Tabrizi SJ, Schobel S, Gantman EC, Mansbach A, Borowsky B, Konstantinova P, Mestre TA, Panagoulias J, Ross CA, Zauderer M, Mullin AP, Romero K, Sivakumaran S, Turner EC, Long JD, Sampaio C; Huntington's Disease Regulatory Science Consortium (HD-RSC). A biological classification of Huntington's disease: the Integrated Staging System. Lancet Neurol. 2022 Jul;21(7):632-644. doi: 10.1016/S1474-4422(22)00120-X. PMID: 35716693).
- Several variables were removed or transformed to minimize participant identification risk (i.e., *height*, *weight*, *maristat*).

Variable changes for PDS6 releases (R1, R2):

Between the minor releases (R1, R2) of the PDS6, there were no changes in the variables (included, excluded or modified). However, certain variable values on the *pharmacotherapy* file were set to missing in the R2 of PDS6, in case *cmtrt__modify* were coded as Investigational drugs to avoid identification. These variables were dose (*cmdostot*), route (*cmroute*), and frequency (*cmdosfrq*).

Number of participants

Enroll-HD is an active, longitudinal study. A participant eligible for inclusion for one PDS release may be ineligible in the next (e.g., participant data quarantined). Table 10 presents sample size by PDS release, while Table 11 and Table 12 summarize the changes in participants between the last version of each PDS and between each release, respectively, regarding the new participants included, participants excluded, and those preserved from the previous release.

Table 10 - Number of participants in each PDS release.

PDS	Release	Number of Participants
PDS1	R1	1457
PDS2	R1	4146
	R2	4146
PDS3	R1	8714
	R2	8714
PDS4	R1	15301
	R2	15301
	R3	15300
PDS5	R1	21116
	R2	21116
PDS6	R1	25550
	R2	25550

Table 11 - Changes in the participants between consecutive PDS releases (final versions).

PDS releases	Changes on the participants from the previous release		
	Same	*Added*	*Removed*
PDS1 → PDS2_R2	1148	2998	309
PDS2_R2 → PDS3_R2	3939	4775	207
PDS3_R2 → PDS4_R3	8470	6830	244
PDS4_R3 → PDS5_R2	15150	5966	150
PDS5_R2 → PDS6_R2	20916	4634	200

Table 12 - Changes in the participants between PDS minor releases.

PDS releases	Changes on the participants from the previous release		
	Same	*Added*	*Removed*
PDS1_R1 → PDS2_R1	1148	2998	309
PDS2_R1 → PDS2_R2	4146	0	0
PDS2_R2 → PDS3_R1	3939	4775	207
PDS3_R1 → PDS3_R2	8714	0	0
PDS3_R2 → PDS4_R1	8470	6831	244
PDS4_R1 → PDS4_R2	15301	0	0
PDS4_R2 → PDS4_R3	15300	0	1
PDS4_R3 → PDS5_R1	15150	5966	150
PDS5_R1 → PDS5_R2	21116	0	0
PDS5_R2 → PDS6_R1	20916	4634	200
PDS6_R1 → PDS6_R2	25550	0	0

Visits

There are several types of visits included in the PDS. All PDS releases include Enroll-HD baseline, follow-up, and phone contact visits. Since PDS3, Enroll-HD unscheduled visits are also in the PDS, alongside baseline and follow-up visits from Registry (R2 and R3), and retrospective visits from ad hoc studies.

Table 13 summarizes the number of visits included in each Periodic Dataset release by visit type. In summary, no changes occurred in visits included between the releases of the same PDS, except for the PDS4_R2 and PDS4_R3, where a baseline visit was removed due to consent withdrawal of one participant, and for the PDS6_R1 and PDS6_R2, where two visits were removed for two family control participants.

Table 13 - Number of visits by type in each PDS release.

	PDS release											
	PDS1	*PDS2*		*PDS3*		*PDS4*			*PDS5*		*PDS6*	
	R1	*R1*	*R2*	*R1*	*R2*	*R1*	*R2*	*R3*	*R1*	*R2*	*R1*	*R2*
Enroll												
Baseline	1457	4146	4146	8714	8714	15301	15301	15300	21116	21116	25550	25550
Follow Up	681	2324	2324	7316	7316	21864	21864	21864	34859	34859	53180	53178
Phone Contact	93	275	275	1114	1114	4008	4008	4008	10917	10917	19528	19528
Unscheduled	-	-	-	43	43	131	131	131	204	204	286	286
Ad-hoc												
Retro-Visits	-	-	-	369	369	809	809	809	970	970	1018	1018
Registry 2												
Baseline	-	-	-	1285	1285	1827	1827	1827	2079	2079	2153	2153
Follow Up	-	-	-	2031	2031	2715	2715	2715	2950	2950	3025	3025
Registry 3												
Baseline	-	-	-	2313	2313	3528	3528	3528	4168	4168	4337	4337
Follow Up	-	-	-	2393	2393	4405	4405	4405	5540	5540	5777	5777
Unscheduled	-	-	-	143	143	204	204	204	232	232	233	233
Total Visits	**2231**	**6745**	**6745**	**25721**	**25721**	**54792**	**54792**	**54791**	**83035**	**83035**	**115087**	**115085**

Coding

Missing Values

There are two overarching categories of missing data in the dataset: system-defined missing data (indicated by blank variable 'entries'), and user-defined missing data (indicated by specific codes, which indicate reason for missingness). These user-defined labels - 'exceptional values' - were introduced in PDS4_R1 and have since remained unchanged. Each one of them is represented in the dataset by a specific code, detailed in Table 14. For more details regarding the missing code values, please consult the **Data Dictionary** document.

Table 14 - Resume of the missing values codes in use since PDS4.

Code	Meaning	Number	Text	Date
Wrong	Value entered but identified as wrong	9996	WRONG	9996-09-09
Not applicable	Value expected, but not entered (does not apply)	9997	NOTAPPL	9997-09-09
Missing	Value expected, but not entered (data collection not performed)	9998	MISSING	9998-09-09
Unknown	Variable response option for a limited set of variables; selected by site when field cannot be completed as unknown to participant.	9999	UNKNOWN	

Medications and Nutritional Supplements

For the PDS, the Enroll-HD participants' use of medications and nutritional supplements are provided in two files: *pharmacotx* and *nutsuppl*, respectively.

The *pharmacotx* file includes coded data from the Enroll-HD database. PDS1 to PDS3 included information from the WHO-DD (WHO Drug Dictionary), which is used to code medications. For PDS4 and later versions, the WHO-DD information is replaced with Enroll-HD specific internal codes for medications in the variable *cmtrt__decod* ("Drug Name – Code"). The internal codes starts with RX, followed by 9 digits. End users with a license for the WHO-Drug Global may request access to the actual WHO-Drug Global codes. Before the release of PDS6, in the EDC the previous version of the WHO-DD (version 2012) had been updated to the WHO-Drug Global (version 2021). Therefore, some of the terms

and codes (*cmtrt__modify, cmtrt__decod, cmtrt__ing, cmtrt__atc)* included in the PDS6 are different from the ones used in previous versions.

The indication for a medication is coded with MedDRA (https://www.meddra.org). From PDS4 on, the indication is coded with the internal Enroll-HD condition code. The internal condition code starts with CX followed by 9 digits. End users with a license for MedDRA may request access to the actual MedDRA codes. Before the release of PDS6 in the Electronic Data Capture (EDC), previous PDS releases used MedDRA (2012 and 2015) versions. Therefore, some of the terms and internal codes values differ between PDS6 and previous versions. For additional details regarding the coding system, please consult the **Coding Systems** document available on the Enroll-HD General Study Document section on https://enroll-hd.org/ .

Comorbidities

For participants in the Enroll-HD Study with documented comorbid conditions, the data are contained in the *comorbid* file of the PDS.

The comorbidities are coded with the ICD-10 code (https://icd.who.int/browse10/2010/en), the variables *mhterm__modify* ("Condition – Modified Term (coded)") and *mhterm__decod* ("Condition – Code (coded)") allow the user to check groups of diseases based on an organ system. Surgeries and procedures are coded using MedDRA. Since PDS4, the Enroll-HD internal condition code is included in the variable *mhterm__decod*. In PDS6 an updated version of MedDRA (2021) was used. Previous PDS releases used version of MedDRA (2015). Therefore, some of the terms and internal codes values differ between PDS6 and previous versions.

End users with a license for MedDRA may request access to the actual MedDRA codes.

For additional details regarding the coding system, please consult the **Coding Systems** document, available on the Enroll-HD General Study Document section on https://enroll-hd.org/ .

Identification risk

To ensure that data included in each PDS release are HIPAA-compliant and that the risk for participant identification is minimized, two methods are employed: 1) "Safe Harbor" method; and 2) "Expert Determination" method. For full details of these methods, please refer to document **Data Quality Management and Participant Privacy**.

Expert Determination recommendations may include the transformation, aggregation, or suppression of certain variables flagged as potentially identifying.

Variables subject to aggregation, and their accompanying thresholds in each PDS release, are described in Table 15 to Table 17.

Pursuant to variable transformation, aggregation, and suppression, data retained in the PDS releases is assessed for individual participant identification risk using the R software package *sdcMicro*. Genotype unknown participants exceeding a 1% identification risk probability are to be excluded from the dataset, for all other participants a 3% identification risk threshold is applied. The identification risk thresholds have remained unchanged across PDS releases. The variables considered in the analysis of identification risk have been updated, as described below:

PDS1 to PDS5 (including all minor releases):

- age (baseline), caghigh, sex, race (aggregated), ISCED (baseline), BMI (baseline).

PDS6:

- age (baseline), caghigh, sex, race (aggregated), ISCED (baseline), BMI (baseline), region.

Aggregation thresholds

To minimize participant identification risk, data aggregation techniques are applied to specific variables for PDS releases. These variables, and the thresholds used for aggregation for each PDS release, might be different: the cohort size and profile allow for such aggregation threshold adjustments while maintaining low identification risk.

Determining a cut-off point to aggregate data is a trade-off between the amount of aggregation and the utility of the data: higher levels of aggregation are associated with lower precision of data.

Table *15* to Table *17* detail the variables aggregated in each PDS and their respective thresholds. All changes are highlighted in yellow. No changes have been made to variable aggregation thresholds since PDS2.

Note: Deaggregated data may be obtained through special request, subject to Scientific Review Committee (SRC) approval. Please refer to the Access Data and Biosamples webpage on **www.enroll-hd.org** for information on how to request a specified dataset (SPS). The **Understand and Interpret the Data** document provides additional detail on handling aggregated values.

Table 15 - Aggregated variables and aggregation thresholds in PDS1.

Data file	Variable	Variable label	Criteria for aggregation
general	*age*	Age at visit	<18 or >79
subject	*caghigh*	Research larger CAG allele determined from DNA	<17 or >55
subject	*caglow*	Research smaller CAG allele determined from DNA	>31
general	*race*	Ethnicity	• **Other**: includes "Native Hawaiian or Other Pacific Islander", "Alaska Native/Inuit", "African - South", "African - North" and "Other"; • **Asian**: includes "Asian – West" and "Asian – East".

Table 16 - Aggregated variables and aggregation thresholds in PDS2.

Data file	Variable	Variable label	Criteria for aggregation
general	*age*	Age at visit	<18
subject	*caghigh*	Research larger CAG allele determined from DNA	>70
subject	*caglow*	Research smaller CAG allele determined from DNA	>28
general	*race*	Ethnicity	• **Other**: includes "Native Hawaiian or Other Pacific Islander", "Alaska Native/Inuit", "African - South", "African - North" and "Other"; • **Asian**: includes "Asian – West" and "Asian – East".

Table 17 - Aggregated variables and aggregation thresholds since PDS3 and subsequent releases.

Data file	Variable	Variable label	Criteria for aggregation
participation	*age_0*	Age at enrollment	<18
adhoc, enroll, registry	*age*	Age at visit	<18
profile	*caghigh*	Research larger CAG allele determined from DNA	>70
profile	*caglow*	Research smaller CAG allele determined from DNA	>28
profile	*race*	Ethnicity	• **Other**: includes "Native Hawaiian or Other Pacific Islander", "Alaska Native/Inuit", "African - South", "African - North" and "Other"; • **Asian**: includes "Asian – West" and "Asian – East".

Quality Control Procedure

Each Enroll-HD PDS goes through stringent Quality Control (QC) procedures prior to release. For full details of these methods, please refer to document **Data Quality Management and Participant Privacy**.

Up to and including PDS5, approximately 200 QC checks were applied to the data. For PDS6, check thresholds were reviewed and updated, and new checks added, for a total of approximately 400. New checks encompassed checks for new variables (e.g., *capscore*; HD-ISS variables) and systematic outlier checks for all continuous variables. More detailed information on each check can be found in the document **Quality Control: Observations and Unusual Findings**.

Participant level values

Changes to individual participant data values between PDS releases, both major and minor, are not described in this document; they are captured and listed in individual spreadsheets, available upon request. These spreadsheets detail individual datum changes and the underlying reasoning for change (e.g., medication codes were changed).

Revision History

Document Name	Summary of Changes
Change Log Overview: PDS1 to PDS6 v20230104	First version for sixth Enroll-HD periodic dataset (PDS6; R1)
Change Log Overview: PDS1 to PDS6 v20230915	Second version for sixth Enroll-HD periodic dataset (PDS6; R2)