Data Cutoff Date – Definition, Algorithm, and Application

1. Objective

The purpose of this guidance document is to provide the definition and algorithm used in determining data cutoff date, and to explain how they are applied in clinical datasets (as specified in Section 3) used for Regulatory submissions, publications, and other analysis purposes.

2. Definition

The data cutoff date (DCO) is defined as the date by which all data records in the clinical datasets will be included in or excluded from a prescribed set of analyses. For the Biologics License Application (BLA) submission, the DCO date is the date of data cutoff date (26-Mar-2021) with exceptions of data related to unblinding, death, serious adverse events (SAE), COVID-19 symptoms, and pregnancy.

A 'target date variable' is defined as the date variable from which the date is to be compared with the data cutoff date. The target date variable in each dataset is to be identified first (as described in Section 3). If the date from the identified target date variable (referred to as the 'target date' thereafter) is on or prior to the cutoff date, then the record containing the target date will be included in the post-cut dataset. If the target date is after the cutoff date, then the record containing the target date will be excluded from the post-cut dataset. In the disposition data set, the study status by DCO date will be reported.

3. Algorithm and Application

The following algorithm and rules are to be followed to identify the target date variable in each dataset to which the cutoff date is to be applied:

3.1 When there is only one date variable (eg, visit date) in the dataset, then that date variable serves as the target date variable, and the target date is to be compared with the cutoff date for all records in that dataset.

3.2 When there is a start date and an end date in a record in the dataset, the start date is used as the target date variable. If the start date is before the cutoff date and the end date is after the cutoff date, then the start date and the end date are kept as is, with the understanding that the entire record (including start and end dates) are cleaned and included in the data version with the specified data cut.

4. Data Handling Convention

When any of the identified target date variables involves partial or missing date, the following conventions are applied in order to determine whether such data records will be included in the post-cut datasets. Partial or missing target date variables are not to be imputed in any post-cut datasets.

Missing day only

When the month and year in the target date variable are the same as or before the month and year in the cutoff date, then the data record containing the partial target date variable with missing day will be included in the post-cut dataset.

Missing month

When the year in the target date variable is the same as or before the year in the cutoff date, then the data record containing the partial target date variable with missing month will be included in the post-cut dataset.

Missing year

When the year in the target date variable is missing, the target date for that data record is considered completely missing. The data record containing the missing year in the target date variable will be included in the post-cut dataset.

5. **Datasets and Date Variables to Be Used for the Data Cutoff Date Algorithm**

The DCO date will be used to cut the source data as listed in Table 1, with exception of the following data where no cut will be done:

- Unblinding
  - Unblinding form in CRF
  - Unblinding date in IRT
  - Date of first dose in open-label phase collected in CRF
  - Treatment information of open label Day 1 in IRT
  - Date of Participant Decision Visit / OL-D1 collected in CRF
- SAE and death
  - Serious AE from AE form in CRF
  - End of study due to SAE or death in CRF
  - Dosing discontinuation due to SAE or death in CRF
- COVID-19
  - Symptom log form in CRF
  - COVID-19 severity form in CRF
  - Adjudication Committee assessments of COVID-19 and severe COVID-19
  - COVID Diagnostic Test form in CRF
  - Elecsys results
  - Immunogenicity Assessment form in CRF
  - Blood Sample Collection for Immunologic Assessment of SARS-CoV-2 Infection form in CRF
  - Results of RT-PCR for SARS-CoV-2 and other pathogens (import LAB)
  - Central Laboratory - Nasopharyngeal Swab form in CRF
  - Central Laboratory - Nasopharyngeal Swab (Single) form in CRF
  - Saliva Collection form in CRF
- Pregnancy

- Pregnancy test form in CRF

Table 1 Datasets and Target Date Variables for the Data Cutoff Date Algorithm

Source	Dataset	Target Date Variable	Variable Label
CRF	AE (where AESER ne Yes and the following variables are equal to 0: AESDTH, AESLIFE, AESHOSP, AESDISAB, AESCONG, AESMIE)	AESTDTTM_RAW	AE start date and time (derived) (Character)
CRF	CB	CBDAT_RAW	Date of assessment (Character)
CRF	CHILLS	CHILLSTMSP_RAW	PC Time Stamp (Character)
CRF	CLABIM_ANT *		
CRF	CM	CMSTDAT_RAW	Start date (Character)
CRF	COVID	COVSTDAT_RAW	Date of missed or out of window visit or assessment (Character)
CRF	CP	CPDAT_RAW	Procedure date (Character)
CRF	CVDSEV *		
CRF	CVCNTCT	SICKDT_RAW	Date of Contact (Character)
CRF	DIARY7	TEMPTMSTP_RAW	PC Time Stamp (Character)
CRF	DIARY7OBS	OBSTMSP_RAW	PC Time stamp (Character)
CRF	DIARY7RXT	DIARY7RXTMSP_RAW	PC Time Stamp (Character)
CRF	DM [1]		
CRF	DS (where DSTERM_STD ne "DEATH" and DSTERM_STD ne "SAE" and DTHDAT is missing)	DSSTDAT_RAW	Date of study discontinuation/completion (Character)
IMPORT	DUKE [2]	COLLECTION_DATE	Sample collection date
IMPORT	DV	Occurred	Occurred
IMPORT	EAS *		
IMPORT	ELECSYS *		
IMPORT	ELEC_BAS *		
CRF	ENROLL	ICDAT_RAW	Date of Informed Consent (Character)
CRF	EOT (where DSTERM_STD ne "DEATH" and DSTERM_STD ne "SAE")	DSSTDAT_RAW	Date of dosing discontinuation (dd MMM yyyy) (Character)
CRF	ESCDIARY	ESCDIARYTMP_RAW	ESCDIARYTMP (Character)

Source	Dataset	Target Date Variable	Variable Label
CRF	EX (where FolderName ne "Participant Decision Visit / OL-D1")	EXSTDAT_RAW	Treatment Date (Character)
CRF	FATIGUE	FATIGUETMSP_RAW	PC Time Stamp (Character)
IMPORT	GCL_TRACKING	COLLECTION_DATE	
CRF	GLAND	GLANDTMSP_RAW	PC Time Stamp (Character)
CRF	HEAD	HEADTMSP_RAW	PC Time Stamp (Character)
CRF	IE [1]		
IMPORT	IG_NAB [2]	COLLECTION_DATE	Sample collection date
IMPORT	IG_PPD	DRAWDATE	Collection Date
CRF	IMM_BLOOD *		
IMPORT	IRT_UNM (where VIS_TYPE doesn't contain "Unblinding" or "Open Label Day 1")	VISIT_DT	System Visit Date/Time
CRF	JOINT	JOINTTMSP_RAW	PC Time Stamp (Character)
IMPORT	LAB *		
CRF	LDT *		
CRF	MH	MHSTDAT_RAW	Start date (Character)
CRF	MUSCLE	MUSCLETMSP_RAW	PC Time Stamp (Character)
CRF	NASAL_SARS *		
CRF	NASAL_SARS2 *		
CRF	NAUSEA	NAUSEATMSP_RAW	PC Time Stamp (Character)
CRF	OTHER	OTHERTMSP_RAW	PC Time Stamp (Character)
CRF	PAIN	PAINTMSP_RAW	PC Time Stamp (Character)
CRF	PE	PEDAT_RAW	Date of examination (Character)
CRF	PT *		
CRF	RAND	RANDDAT_RAW	Date of Randomization (Character)
CRF	RED	REDTMSP_RAW	PC Time Stamp (Character)
CRF	RISK [1]		
CRF	SALIVA_LOG *		
CRF	SC	SCCONDAT_RAW	Date of Contact or Contact Attempt (Character)
CRF	SRF	SUB_DATE_RAW	Sub_date (Character)
CRF	SUBJECT [1]		
CRF	SWELL	SWELLTMSP_RAW	PC Time Stamp (Character)
CRF	SYMPLOG *		
CRF	UNBLND *		
CRF	UNS	VISITDAT_RAW	Visit Date (Character)
CRF	VISIT (where FolderName ne	VISITDAT_RAW	Visit date (Character)

Source	Dataset	Target Date Variable	Variable Label
	"Participant Decision Visit / OL-D1")		
IMPORT	VRC [2]	COLLECTION_DATE	Sample collection date
CRF	VS	VSDAT_RAW	Date of Measurements (Character)
CRF	VSDOSE	VSDAT_RAW	Date of Measurements (Character)

* exception for data cutoff.

[1] Dataset doesn't have a target variable.

[2] Dataset should be merged with GCL_TRACKING data by Subject and accession number to populate Collection_date.

Table 2 Datasets and Target Date Variables for Data Cutoff Exception

Source	Dataset	Target Date Variable	Variable Label
CRF	AE (where AESER eq "Yes" or at least one the following variables is equal to 1: AESDTH, AESLIFE, AESHOSP, AESDISAB, AESCONG, AESMIE)	AESTDTTM_RAW	AE start date and time (derived) (Character)
CRF	CLABIM_ANT	LBDAT_RAW	Collection date (Character)
CRF	CVDSEV	ARDSSTDT_RAW	ARDS Start Date (Character)
		CLINSTDT_RAW	Clinical Evidence Date of Assessment (Character)
		ECMOSTDT_RAW	ECMO Start Date (Character)
		HEPSTDT_RAW	Hepatic Dysfunction Start Date (Character)
		HRSTDAT_RAW	Heart Rate Start Date (Character)
		ICUSTDT_RAW	ICU Start Date (Character)
		MECHSTDT_RAW	Mechanical Ventilation Start Date (Character)
		MMHGSTDT_RAW	MMHG Start Date (Character)
		NEURSTDT_RAW	Neurologic Dysfunction Start Date (Character)
		NINVSTDT_RAW	Non-Invasive Ventilation Start Date (Character)
		OXSTDT_RAW	High-Flow Oxygen Start Date (Character)
		PAO2STDAT_RAW	PAO2 Start Date (Character)
		RADSTDT_RAW	Radiographical Date of Assessment (Character)
		RENSTDT_RAW	Acute Renal Dysfunction Start Date (Character)
		RESPSTDT_RAW	Respiratory Failure Start Date (Character)
		RSPSTDAT_RAW	Respiratory Rate Start Date (Character)

Source	Dataset	Target Date Variable	Variable Label
		SPO2STDAT_RAW	SPO2 Start Date (Character)
		VASOSTDT_RAW	Vasopressors Start Date (Character)
CRF	DS (where DSTERM_STD eq "DEATH" or DSTERM_STD eq "SAE" or DTHDAT is not missing)	DSSTDAT_RAW	Date of study discontinuation/completion (Character)
IMPORT	EAS	COVIDT	COVID-19 AC-ASSESSED ONSET DATE: DD-MMM-YYYY
		SVCOVIDT	SEVERE COVID-19 AC-ASSESSED ONSET DATE: DD-MMM-YYYY
IMPORT	ELECSYS	DRAWDATE	Collection Date
IMPORT	ELEC_BAS	DRAWDATE	DRAWDATE
CRF	EOT (where DSTERM_STD eq "DEATH" or DSTERM_STD eq "SAE")	DSSTDAT_RAW	Date of dosing discontinuation (dd MMM yyyy) (Character)
CRF	EX (where FolderName eq "Participant Decision Visit / OL-D1")	EXSTDAT_RAW	Treatment Date (Character)
CRF	IMM_BLOOD	IMMDT_RAW	Date of Collection (Character)
IMPORT	IRT_UNM (where VIS_TYPE contains "Unblinding" or "Open Label Day 1")	VISIT_DT	System Visit Date/Time
IMPORT	LAB	LBDTM	Actual Collection Date and Time
CRF	LDT	LDTDAT_RAW	Date of Test (Character)
CRF	NASAL_SARS	LBDAT_RAW	Collection Date (Character)
CRF	NASAL_SARS2	LBDAT_RAW	Collection Date (Character)
CRF	PT	PTDAT_RAW	Date of test (Character)
		LBDAT5_RAW	Collection date (Character)
CRF	SALIVA_LOG	SALIVADT_RAW	Date of Collection (Character)
CRF	SYMPLOG	SYMPDAT_RAW	Date (Character)
CRF	UNBLND	UNBLNDDAT_RAW	Date of unblinding (Character)
		UNBLND_IFCDAT_RAW	Date of updated informed consent (Character)
CRF	VISIT	VISITDAT_RAW	Visit date (Character)