Data Cut-off Approach for SDTM Data

Monitha M Harilkumar,
Natalia Smeljanski

Merck

Data Cut-off Approach for SDTM
Agenda

- **General**
  - ➢ Data cut-off –Objective and Scope
  - ➢ Points to consider – 5 Hows
  - ➢ „Terms" – Glossary for term definition

- **Domain specific instructions**
  1. Special purpose domains
  2. Interventions
  3. Events
  4. Findings

- **Workflow**

Merck

General
Data Cut-off : Objective and Scope

Objective:

To provide guidance and instructions on the implemementation of data cut-off in order to:

- Minimize gaps in Standards compliance during cut-off implementation
- Ensure valid SDTM structure and logical relationships of data in different domains
- Establish an internal guideline/documentation to facilitate a streamlined process

This objective required input from both statisticians and SDTM experts

Scope today:

To describe the SDTM-cut-off approach developed based on lessons learned and experiences from previous submissions. ADAM datasets are out of scope.

Merck

What is Data Cut-off and how does it look in SDTM?

Data-cut-off: A process, where clinical data is cut into subsets.
The cut is based on the rule set, which should be described for the reviewer in the reviewer's guides (e.g. cSDRG or ADRG).
Data-cut-off could be executed for i.e. **DSUR**, Interim Analysis, Monitoring committee, or regulatory requests.

Ideal SDTM form at the end of the study
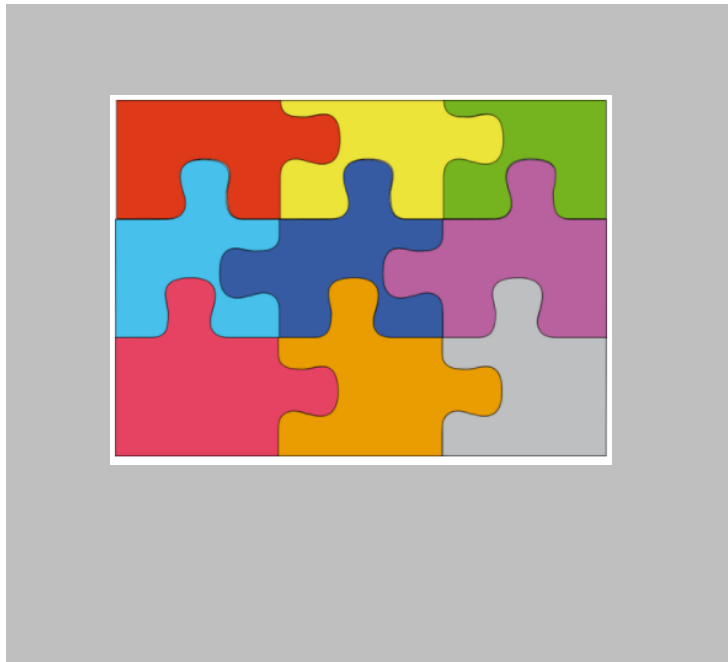
Data split in different domains, structuraly divided, however logically connected and complete:

Examples:

- Tumor data

- PK data

- Main and supplemental domains

- DM and Exposure

Data cut-off aproach for SDTM /2017

Data cut-off :Real vs Ideal world

Data cut on SDTM,
ideal expectation

Data cut on SDTM,
real situation

Points to consider for Data-Cut off: The 5 Hows

- **How to proceed first –** **T**o select data sub-set and derive after, or derive first and select data sub-set after

- **How to keep –**
- **L**inked information in related domains (e.g. PC-PP; TU-TR-RS)

How to approach - **I**mportant data (safety data, protocol deviations)

- **How to handle -** **I**ncomplete dates

- **How to cut -** **S**upplemental data set (no dates available)

Data cut-off approach for SDTM /2017

Some terms and definitions

➢ **Cut-off subject:** A cut-off subject refers to a subject on which cut-off rules are applied.

➢ **Treated subject:** A "Treated subject" in terms of the cut-off is defined as a subject where RFSTDTC is not blank and who has at last one exposure record.

➢ **[PLUS] datasets:** Datasets containing the parent SDTM data merged with their supplemental variables and consists of comments (from Comments Domain) where a reference to a record in a parent dataset exists. These datasets are indicated as [PLUS] datasets.

Note: The cut-off program is done on [PLUS] datasets. No cut-off will be applied specifically on dates in supplementary domains.

MERCK

Domain specific instructions
1. Special Purpose Domains

> ## <u>Demographics (DM)</u>

RFICDTC (Informed consent date) is used as reference date for all subjects. Subjects that start treatment after cut-off, but have signed informed consent before will stay in the data.

<u>Step 1: Selection of subjects to be included in the Cut-off DM [PLUS] Dataset</u>
Include all subjects where RFICDTC <= cut-off date

<u>Note:</u>

Subjects that start treatment after cut-off, but have a signed informed consent before or on cut-off date will stay in the data.

Will take SCRNFAIL, NOTASSGN and NOTTRT into account and include in DM as well

MERCK

Special Purpose Domains

Selection of data in DM domain-Example

CUT OFF date: 05 Jan 2020

USUBJID	ICDTC	RFSTDTC	ACTARM
S001	01.01.2020	05.01.2020	TRT1
S002	05.01.2020	14.01.2020	TRT1
S003	10.01.2020	15.01.2020	TRT1
S004	02.01.2020		SCRNFAIL
S005	05.01.2020		NOTASSGN
S006	15.01.2020		NOTASSGN

Note: RFSTDTC is derived after cut-off as per our preferred approach.

As per our approach , we will consider records of SCRNFAIL/NOTASSIGN based on RFICDTC.

Full data set

(will be delivered for safety and DSUR reports)

Although RFSTDTC date falls after cut-off , they will be included in DM but not in EC/EX domain

Note: however ACTARM will not change

Special Purpose Domains
<u>Derivation of reference variables in Dataset DM [PLUS]</u>

<u>Step 2: Reference variables in Dataset DM [PLUS]</u>

Preferred approach is that the derivation of variables like RFSTDTC, RFENDTC, RFXSTDTC, RFXENDTC, RFPEXNDTC, DTHDTC, DTHFL will be performed after cut-off

- RFPENDTC: is derived based on information from the study termination page or the death date (whichever is the final record). If the RFPENDTC is blank for a cut-off subject, then the RFPENDTC will be set to the data cut-off date.
- **An alternative option** is to derive the reference variables first and then to proceed with data cut-off.

<u>Note:</u>

To decide earlier of preferred approach or alternative option to be applied before cut-off.

2. Interventions
Exposure Datasets- EX and EC

Cut-off rule should be applied to all interventions domains for those subjects that remain in DM after applying the cut-off rule based on RFICDTC.

Any records where –STDTC > cut-off date will be excluded and no derivation of variables like-ENTPT.

CUT OFF date: 05 Jan 2020

DM:USUBJID	DM.RFICDTC	DM.RFSTDTC	DM:ACTARM
S001	01.01.2020	05.01.2020	TRT1
S002	05.01.2020	14.01.2020	TRT1

EC.USUBJID	ECTRT	ECSTDTC	ECENDTC
S001	TRT1	05.01.2020	06.01.2020
S002	TRT1	14.01.2020	16.01.2020

Subjects will not have data in EX domain, however ACTARM will not change.

Note: RFSTDTC is derived after cut-off as per our preferred approach and therefore, we will consider records of SCRNFAIL/NOTASSIGN regardless of RFICDTC.

MERCK

2. Events Class
Adverse Events Domain

Cut-off rule should be applied to all events domains for those subjects that remain in DM after applying the cut-off rule based on RFICDTC.

Adverse Events (AE)

- --STDTC: Any records where --STDTC > cut-off date will be excluded

- --ENDTC and --ENTPT: No changes due to cut-off should be made on these variables.

Cut off date = 01Jan2021
USUBJID =001
AESTDTC= 30 Dec 2020
AE ongoing
End Time Point: end of study date
End of study = 30 Jan 2021

After Cut-off: This record will be included, as AESTDTC is before cut-off date.

Cut off date = 01Jan2021
USUBJID =001
AESTDTC= 02 JAN 2021
AE ongoing
End Time Point: end of study date
End of study = 30 Jan 2021

After Cut-off: This record will not be included, as AESTDTC is after cut-off date.

2. Events Class

Protocol Deviation (DV)

Include **ALL** protocol deviations for those subjects which are selected for DM (after applying cut-off rule on Informed Consent Date).

Exception: Protocol deviation date (DVDTC) is usually a date when deviation was collected and not when the protocol deviation occurred. But if DVDTC contains the real start date of a protocol deviation (e.g. for derived DVs) then the cut-off date shall be applied on DVDTC.

MERCK

2. Events Class
Protocol Deviation (DV) Example

Cut off Date – 05JAN 2020

Cut-off based on RFICDTC

Cut-off based on DVDTC

DM				DV		
USUBJID	ICDTC	RFSTDTC	ACTARM	USUBJID	Protocol deviation	Deviation date
S001	01.01.2020	05.01.2020	TRT1	S001	Important dev	04.01.2020
S001	01.01.2020	05.01.2020	TRT1	S001	Clinically important dev	10.01.2020
S001	01.01.2020	05.01.2020	TRT1	S001	Not important	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Important dev	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Clinically important dev	06.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Not important	07.01.2020
S003	10.01.2020	15.01.2020	TRT1	S003	Important dev	20.01.2020
S003	10.01.2020	15.01.2020	TRT1	S003	Clinically important dev	21.01.2020
S003	10.01.2020	15.01.2020	TRT1	S003	Not important	21.01.2020

DM				DV		
USUBJID	ICDTC	RFSTDTC	ACTARM	USUBJID	Protocol deviation	Deviation date
S001	01.01.2020	05.01.2020	TRT1	S001	Important dev	04.01.2020
S001	01.01.2020	05.01.2020	TRT1	S001	Clinically important dev	10.01.2020
S001	01.01.2020	05.01.2020	TRT1	S001	Not important	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Important dev	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Clinically important dev	06.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Not important	07.01.2020
				S003	Important dev	20.01.2020
				S003	Clinically important dev	21.01.2020
				S003	Not important	21.01.2020

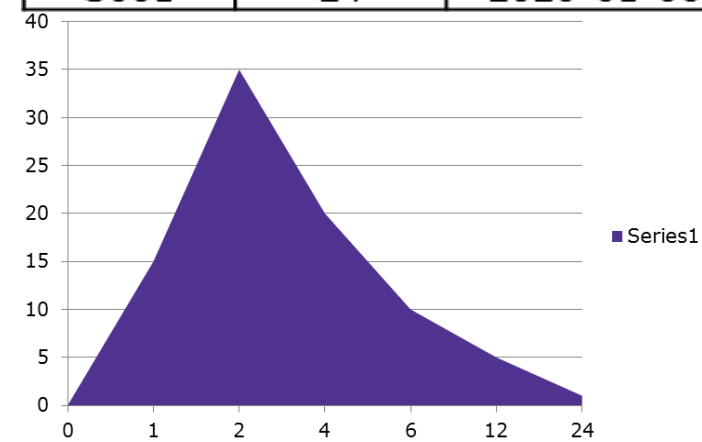| DM				DV		
USUBJID	ICDTC	RFSTDTC	ACTARM	USUBJID	Protocol deviation	Deviation date
S001	01.01.2020	05.01.2020	TRT1	S001	Important dev	04.01.2020
S001	01.01.2020	05.01.2020	TRT1			
S001	01.01.2020	05.01.2020	TRT1	S001	Not important	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Important dev	05.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Clinically important dev	06.01.2020
S002	01.01.2020	06.01.2020	TRT1	S002	Not important	07.01.2020

3. Findings Class
Pharmacokinetics Domains

Keep the complete PK profile. If at least one concentration time point will be before or on date of cut-off, all concentrations of this profile will stay in the data. Example: If "0 hour post dose" and "1 hour post dose" are before or on the cut-off date then all timepoints that are followed by such "24 hours post dose", will still stay in the data.

USUBJID	PCTPT	PCDTC	PCORRES
S001	0	2020-01-05T13:30	0
S001	1	2020-01-05T14:30	15
S001	2	2020-01-05T15:30	35
S001	4	2020-01-05T17:30	20
S001	6	2020-01-05T19:30	10
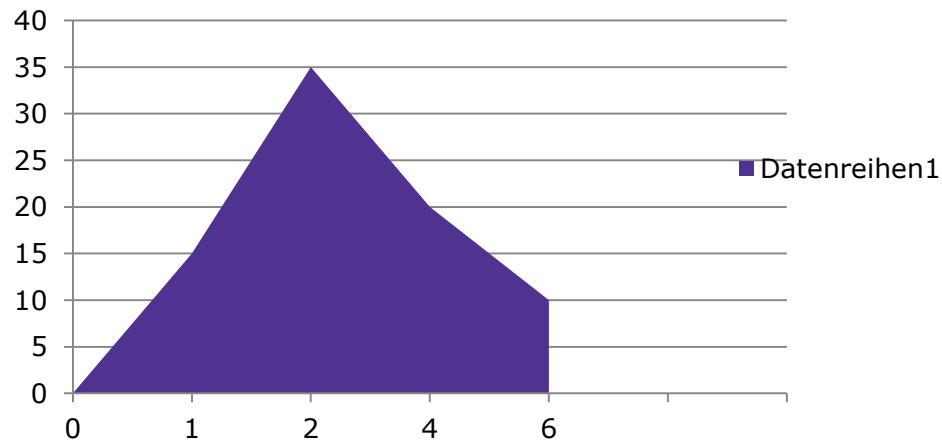
USUBJID	PCTPT	PCDTC	PCORRES
S001	0	2020-01-05T13:30	0
S001	1	2020-01-05T14:30	15
S001	2	2020-01-05T15:30	35
S001	4	2020-01-05T17:30	20
S001	6	2020-01-05T19:30	10
S001	12	2020-01-06T01:30	5
S001	24	2020-01-06T13:30	1

3. Data Cut-off Approach for SDTM
Workflow

[PLUS] data creation-Full data set
- Merge main and supplemental data sets
- Or work on raw data

DM
Create CUT-OFF sub-set of subjects based on IC date

Cut
Events (consider exemptions)
Interventions
Findings (consider relations)

Subject Visits and Elements
- SV
- Apply SV Cut considering all observations collected at correspondent visit
- SE
- Create SE based on SV cut-off result

Derive variables
- Derive correspondent variables in DM (DMRFDTC, DMXSTDTC etc.)
- Derive required variables in all further domains (e.g. –DY, -STDY)

MERCK

Data Cut-off Approach for SDTM
Summary

1. Consider the **5- Hows**
2. Discuss approach in detail with Biostatistics (and any further potential data customer)
3. Create a workflow
4. Describe detailed approach for reviewer (e.g. CDRG)

Cut-off is not a straight forward process, in some cases it needs project-specific decisions!

MERCK

THANK YOU!

MERCK

BACK-up

Merck

Special Purpose domains: SV domain

Cut-off should be applied on SVSTDTC for those subjects selected for DM (after applying cut-off rule on Informed Consent Date). I.e. records with SVSTDTC > Data-cut-off date will be excluded for subjects remaining in DM after cut-off.

Example: data cut-off is 2015-07-31 and the CRF collected data are as follows:
Date of treatment termination visit **2015-08-01**
Date of ECG data collection at treatment termination **2015-07-30**
Date of Lab data collection at treatment termination **2015-08-01**

SDTM data before data cut-off filtering:

SV domain (subject visits): observation with VISITNUM=1000 (treatment termination) and SVDTC (date) **2015-08-01**
EG domain (ECG): observation with VISITNUM=1000 (treatment termination) and EGDTC (date) **2015-07-30**
LB domain (LAB data): observation with VISITNUM=1000 (treatment termination) and LBDTC (date) **2015-08-01**

Preferred Approach :
- **SV** domain (subject visits) with SVSTDTC (date) **2015-07-30** stays (to get consistent link between domains and avoid validation issue in validation report)
- **EG** domain (ECG) with EGDTC (date) **2015-07-30** observation stays (as date of assessment before date of cut-off)
- **LB** domain with LBDTC (date) **2015-08-01** observation will be deleted (as date of assessment after date of cut-off)

Findings Domain: Tumor domains

Cut-off rule should be applied to all findings domains for those subjects that remain in DM after applying the cut-off rule based on RFICDTC. All records in findings domains, where --STDTC > cut-off date, will be excluded.

Tumor domains: Tumor Identification (TU), Tumor Results (TR), Disease Response (RS)

Cut-off to be applied on the response assessment date.

Date of response is the earliest date of the assessed tumor measurements, unless or otherwise specified in SAP. All tumor measurements linked to a response assessment (at one specific time point) should stay as it is and are not applicable to data cut-off.

Findings Domain: Tumor domains : Example

Cut-off date 1 Jan 2020

Assessment visit duration app 3 days.

Target Tumor 1 measured at 30 Dec 2019 and shown SD

Target Tumor 2 measured at 2 Jan 2020 and shown PD.

Target response assessment based on measurements of the same assessment visit is PD.

Case a: Date of target response (PD) set up as latest date of measurement -**2 Jan 2020** .
After Cut-off: all related and linked observations will not be included as response assessment is after cut-off

Case b: Date of target response (PD) set up as earliest date of measurement -30 Dec 2019
After Cut-off: all related and linked observations will stay in data base because response assessment is before cut-off—
Preferred Approach