PharmaSUG China 2023 - Paper CC-137

Implementation of Cut-Off Date in Analysis of Solid tumor Clinical Trials

Amelia Pang, Zhaolei Lv, Sanofi Corporation

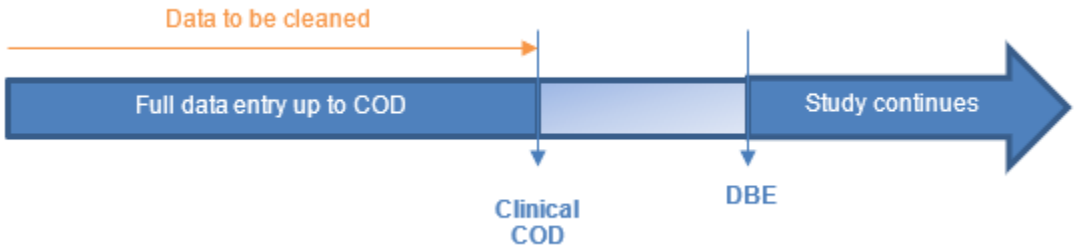
ABSTRACT

In an on-going clinical trial with event endpoints, between the real occurrence of event and the database reflection, usually there is a time delay. It's due to the dynamic data refreshing and the static data extraction on the cut-off date. As IA submission was accepted by HA increasingly, oncology therapeutic area creates more business needs in programming cut-off on late phase study with on-going data. This paper describes the cut off rules and reason in Sanofi with solid tumor indication and focus on How to select participants; How to keep linked information in related domains (TU-TR-RS, EC-EX). How to deal with the records without dates or with partial dates; especially How to approach post-COD records. This paper will also introduce a SAS macro which manage the cut-off based on raw data.

INTRODUCTION

In oncology clinical trials, primary endpoint is usually specified as a time-to-event endpoint such as Progression-Free-Survival (PFS) or Overall Survival (OS). In this context, the statistical analyses are triggered by a pre-specified number of events, which is observed at a date referred as the "cutoff date" (COD). However, there is often a gap between the targeted event occurrence and its knowledge. As the operational COD is based on a planned/estimated COD. The actual COD is used in statistical analyses. Additionally, the data collection, which is based on electronic data capture (EDC) system, is still open for sites to enter or correct new data between the COD and the data extraction; therefore, some of these data may have been collected after the COD. The data cutoff process is then used to represent the status of the data at the COD.

Figure 1



The objective of using a standard macro program to manage the data cutoff is to ensure a consistent approach across Sanofi projects with solid tumor indication. This standard macro program aims at selecting a subset of the data so that the datasets contain only data up to and including the COD.

The scope of use of this standard macro program is limited to randomized phase 2 and 3 clinical trials with a time-to-event endpoint, for the final analysis and any interim analysis (IA) assessing efficacy. All analyses (efficacy and safety) will be performed on the same database, ie data selected after application of this standard macro. The use of this standard macro program is not applicable to phase 1 clinical trials, phase 2 clinical trials with a primary endpoint not specified as a time-to-event, safety review, DMC and IA for futility.

HOW TO SELECT PARTICIPANTS

Informed consent(s) are given by a participant when they allow their data to be used as part of a clinical study. A signed informed consent form is required by all participants to enter a clinical study. If a participant has not given his/her consent by the COD, then the participant and any records related to that participant should be removed from all data provided for the analysis. An informed consent date is not always the same as the screening date. A participant may provide his/her informed consent prior to the COD and then attend a screening visit after. In relation to applying the COD logic based on informed consent, only the date of consent is considered. In case several informed consent forms are to be signed by participant, this is the first IFC signed that is to be considered.

HOW TO KEEP LINKED INFORMATION IN RELATED DOMAINS

There are always relationships between two (or more) datasets, like EX-EC, TR-TU-RS, PC-PP, etc. All relationships make use of the standard domain identifiers, STUDYID, USUBJID. In addition, --REFID, --GRPID, --LNKID can also be used to capture a sponsor-defined or external identifier. And the datasets relationships will be represented in RELREC (related records) dataset. The relationship should be considered in COD logic, keep the consistence between datasets and consistence between analysis reports.

TR-TU-RS

The TU domain represents data that uniquely identifies tumors/lesions (i.e., malignant tumors, culprit lesions, and other sites of disease, e.g., lymph nodes). TR domain represents quantitative measurements and/or qualitative assessments of the tumors or lesions identified in the tumor/lesion identification (TU) domain. These two domains each have a distinct purpose and are related to each other, and may also be related to assessments in the Disease Response and Clin Classification (RS) domain - for the assessment of disease response to therapy. A same cut off logic should be applied.

The data cut strategy will be based on the cycle start date, all the records in a cycle which start date is before COD will be kept. Because: 1. Tumor assessments are collected by cycle. Tumor assessments not always in a same day in a cycle. 2. It is based on the completed information in a cycle when we calculated the target lesion response and best overall response using RECIST criteria.

EX-EC

Collected exposure data points are to be represented in the EC domain. And the EX domain is recognized in most cases as a derived dataset where EXDOSU reflects the protocol specified unit per study treatment. EX and EC will use a same cut off logic.

Exposure data are collected by cycle, same as TR-TU-RS, the data cut strategy will be based on the cycle start date. All exposure records from cycles with a start date equal to or lower than the COD will be kept. Otherwise, if the cycle start date is greater than the COD, the cycle exposure records will be removed.

PC-PP can be also cut use this method. If there is no date information or visit information in one dataset (have related datasets), we can use –REFID/GRPID/--LNKID to select the records which consistent to it's related dataset.

HOW TO DEAL WITH THE RECORDS WITHOUT DATES OR WITH PARTIAL DATES

General rules for date partial missing or fully missing compared with COD is that: If date is fully missing or only day is known, data is kept; if only month and year and they are ≤ month and year of COD, data is kept; if only year and year is ≤ year of COD, data is kept. In other words, include records in the data cut unless the relevant date is after the clinical cutoff date.

HOW TO APPROACH POST-COD RECORDS

Some records collected after the COD could be kept in the database. Such as tumor assessment data and exposure data and all the death/ subject status and Last contact date!

TUMOR EVALUATION DATA

For the tumor evaluation data based on investigator/local radiologists' assessment. Indeed, for a single tumor evaluation, several assessments might be performed at different dates (e.g., several target lesions assessment, non-target lesions, new lesions). Therefore, all records from a single tumor evaluation will be kept if the first assessment belonging to the respective tumor evaluation is performed before the COD.

For tumor evaluation data based on independent blinded review committee (IRC) assessment, all records with an assessment date equal to or before COD based on the Overall Response date will be kept.

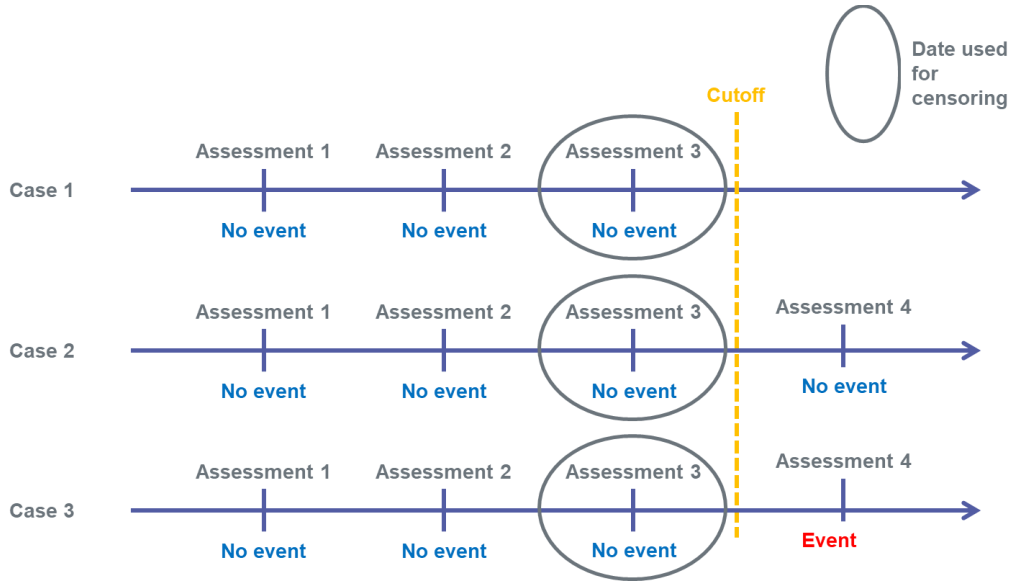
EXPOSURE DATA

Exposure data are collected by cycle, therefore the data cut strategy will be based on the cycle start date. All exposure records from cycles with a start date equal to or lower than the COD will be kept. Otherwise, if the cycle start date is greater than the COD, the cycle exposure records will be removed.

ENDPOINT EVALUATION BASED ON ASSESSMENT

When the primary endpoint event occurrence is based on an assessment, such as a tumor assessment for the PFS, a patient without an event observed before the analysis COD will be censored at the date of the last valid disease assessment before the COD. Data collected after the COD will not be used to censor the patient. Illustration is provided in Figure 2. Therefore, for this endpoint, it may not be necessary to keep post-COD data to justify the specification of the censoring date.

Figure 2 – Illustration of censoring scheme when the primary endpoint event occurrence is based on an assessment

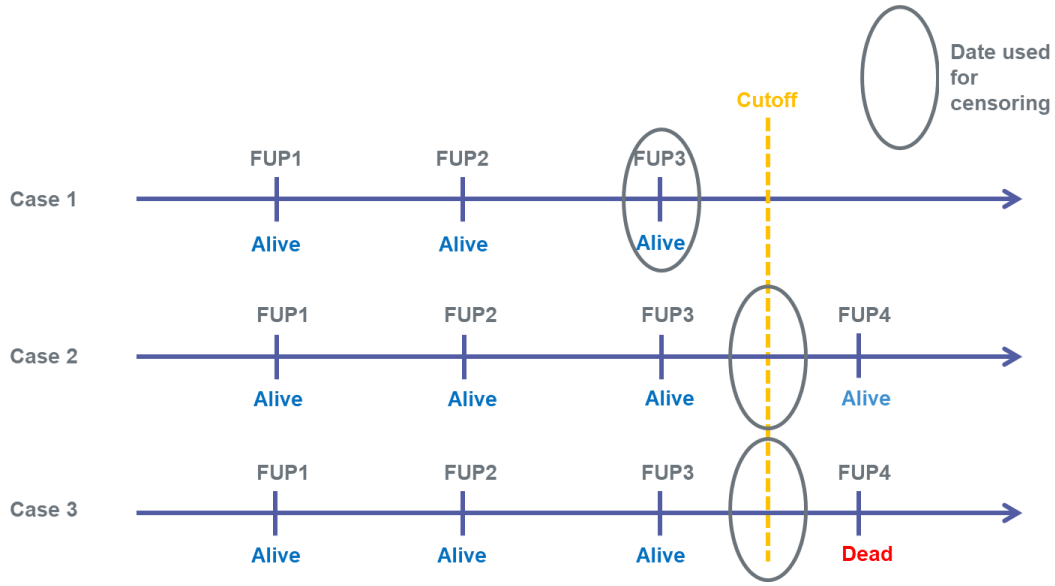


ENDPOINT EVALUATION BASED ON SURVIVAL

When the primary endpoint event occurrence is based on survival, such as OS, a patient without an event observed before the analysis COD will be censored at the last date the patient was known to be alive. Data collected after the COD can be used to censor the patient at the COD. Illustration is provided in Figure 3. Therefore, for this endpoint, it may be necessary to keep post-COD data to justify the specification of the censoring date.

Figure 3 – Illustration of censoring scheme when the primary endpoint event occurrence

is based on survival



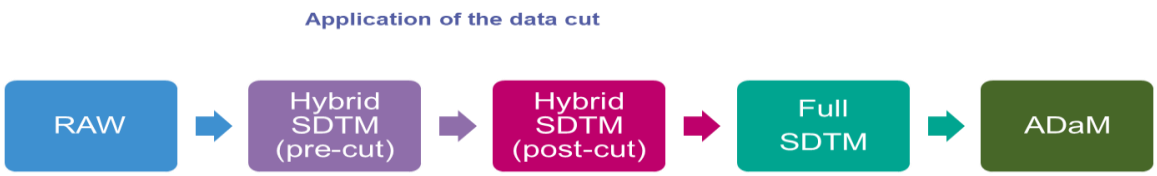
CUT OFF MACRO PROCESS

DATA CUT STRATEGY

The data cut strategy determines the rules to select the subset of data to be included in the analysis.

The data cut is applied to hybrid SDTMs. Full SDTMs and ADaMs datasets are created from hybrid SDTMs post-cut and therefore, will only contained the subset of data used for the analysis.

Figure 4 - Application of the data cut



MANAGEMENT OF CUT-OFF DATA

As consider into all the specific cases above, the final macro is developed by below 3 steps:

1. Keep all records,
2. Keep records which start date or collect date before cut-off date (COD), with –DTC/--STDTC <= COD,
3. Keep all the records in a cycle which start date is before COD. For this, we firstly identify the last cycle (start date) before COD. Then we will select records by cycles which <= last cycle.

The macro can be called:

```
%cut_off(list_sdtm_rule1 =
        ,list_sdtm_rule2 =
```

```
        ,list_sdtm_rule3 =
);
```

CONCLUSION

COD implementation can potentially have a major impact on the trial results. It's important to understand COD principles, develop specifications and programs/macros to apply the COD algorithms correctly and accurately. Data traceability and integrity and include as much information as possible are major considerations in COD applications. Defining and implementing a standard process that includes developing standard programs/methods could increase efficiency and reduce variability across studies. In addition, the operational challenges should not be overlooked. A coordinated effort across functional teams to ensure data quality is a MUST before applying COD analysis. Although oncology COD principles are used to illustrate COD applications in this paper, it can be expanded to any other therapeutic areas or studies where COD is applicable.

REFERENCES

Mei Dey, AstraZeneca "Implementation of Data Cut Off in Analysis of Clinical Trials" available at PharmaSUG-2018-DS19.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Amelia Pang

E-mail: Amelia.Pang@Sanofi.com

Zhaolei Lv

E-mail: Zhaolei.Lv@Sanofi.com