Prepare ADaM data sets for Objective Response Rate (ORR) and Progression-Free Survival (PFS) analysis in an efficient way

Sen Fan, CSPC Pharmaceutical Group Ltd

ABSTRACT

Objective Response Rate (ORR) and Progression-free survival (PFS) are commonly used as efficacy endpoints in solid tumor clinical trials. However, for inexperienced statistical programmers, Best Overall Response (BOR) confirmation rules and PFS censoring schemes are quite obscure. When programmers try to implement these derivations, they may find it hard to ensure the conditions/logics are mutually exclusive and collectively exhaustive. This paper is aimed to help programmers to build up basic knowledge on BOR and PFS. Additionally, with using mockup data as examples, this paper demonstrates how to efficiently prepare ADaM data sets for ORR and PFS analysis.

HOW TO DERIVE BOR WHEN CONFIRMATION OF CR/PR IS REQUIRED

ORR, defined as the proportion of patients with a BOR of Complete Response (CR) or Partial Response (PR) to treatment, is an endpoint commonly used in solid tumor trials. In trials where confirmation of CR/PR is not required, the BOR derivation is straightforward -- just to select the best response (CR>PR>SD>PD>NE) across all time point assessments. Whereas, in non-randomised trials where response is the primary endpoint, to avoid potential bias, confirmation of CR/PR is generally required. The interpretations on CR/PR confirmation rules are given in the <New Response Evaluation Criteria in Solid Tumors guideline> (RECISIT V1.1), as shown below Table 1. However, for statistical programmers who are new to oncology studies, it is hard to understand these rules, and especially, to implement these rules by programming.

Overall response First time point	Overall response Subsequent time point	BEST overall response
CR	CR	CR
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.
a If a CR is *truly* met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

Table 1 Best overall response when confirmation of CR and PR required.

From RECISIT V1.1 guideline, there are some points to be noticed before you derive confirmed BOR.

1. SD should not follow a CR. As explained by the footnote of Table 1, when there is such data, evaluators need to confirm whether response for the second time point should be actually judged as PD or response for the first time point should have been judged as PR instead of CR.

2. Also, generally PR should not follow a CR. However, a rare case can be an exception. If target response is CR and non-target response is NE and without any new lesions at a time point, the corresponding overall response will be PR. PR, in this case, can follow a CR.

3. CR/PR can be claimed only if the criteria are met again at a subsequent time point, and time interval between the subsequent and initial CR/PR is no less than a protocol specified minimum window (e.g., 4 weeks).

4. When SD is believed to be BOR, it must meet a protocol specified minimum time (e.g., 6 weeks) from baseline.

5. When no imaging/measurement is done or only a subset of lesions are measured at a particular time point, the response could be classified as "Not Evaluable (NE)". NE normally will not add any value in the derivation of BOR if there are any other adequate responses. It is noted in the RECISIT V1.1, "In most trials, it is reasonable to consider a patient with time point responses of PR-NE-PR as a confirmed response."

6. According to RECISIT V1.1, "Once a PD, always PD". Normally, study protocol pre-specify that once PD has been documented for a patient, tumor status will not be followed up any more. We should aware that, any tumor responses collected after the first documented PD date, even if a CR or PR, should not contribute to the BOR determination.

7. Some intercurrent events (ICEs), including that patient discontinues the investigational treatment due to adverse events/lack of efficacy or patient starts a new anti-cancer therapy, may affect the existence or interpretation of outcome data. Normally, study protocol/SAP will address on whether to consider the tumor assessments which are collected after ICEs. Sometimes, sensitivity analysis with using a different approach from the primary analysis may also be conducted.

Therefore, we can translate the RECISIT V1.1 confirmed BOR rules into below logics.

With using tumor responses across all time points which are recorded from start date (e.g., randomization date) until disease progression or any pre-specified ICE (e.g., start date of new anti-cancer therapy), BOR for patients are derived in below order:

CR: Patient had CR at an initial time point and had CR at a subsequent time point which is >= 4 weeks apart from the initial one. There can be no other responses except CR/NE between the initial and subsequent time points;

PR: Patient had PR at an initial time point and CR/PR at a subsequent time point which is >= 4 weeks apart from the initial one. There can be no other responses except CR/PR/NE between the initial and subsequent time points;

SD: Patient at least had one response assessment of CR/PR/SD, which is >= 6 weeks apart from the start date;

PD: Patient had PD;

NE: Patient did not meet any of the above criteria.

To make it clear and concise, a mockup data set is created. All typical BOR cases are included in the WROK.RS0, as shown in below Table 2.

USUBJID	RANDDT	NCTXSDT	RSSEQ	RSTESTCD	RSTEST	RSSTRESC	RSDTC
Subject ID	Date of Randomization	Start Date of New Anti-cancer Therapy	Sequence Number	Response Assessment Short Name	Response Assessment Name	Response Assessment Result in Std Format	Date/Time of Response Assessment
A1	2021-12-31		1	OVRLRESP	Overall Response	PR	2022-02-09
A1	2021-12-31		2	OVRLRESP	Overall Response	CR	2022-03-22
A1	2021-12-31		3	OVRLRESP	Overall Response	CR	2022-05-06
A2	2021-12-31		1	OVRLRESP	Overall Response	PR	2022-02-09
A2	2021-12-31		2	OVRLRESP	Overall Response	CR	2022-03-22
A2	2021-12-31		3	OVRLRESP	Overall Response	NE	2022-05-06
A2	2021-12-31		4	OVRLRESP	Overall Response	CR	2022-06-18
A3	2021-12-31		1	OVRLRESP	Overall Response	PR	2022-02-09
A3	2021-12-31		2	OVRLRESP	Overall Response	CR	2022-03-22
A3	2021-12-31		3	OVRLRESP	Overall Response	CR	2022-04-04
A3	2021-12-31		4	OVRLRESP	Overall Response	CR	2022-05-06
B1	2021-09-17		1	OVRLRESP	Overall Response	SD	2021-10-27
B1	2021-09-17		2	OVRLRESP	Overall Response	PR	2021-12-07
B1	2021-09-17		3	OVRLRESP	Overall Response	PR	2022-01-17
B1	2021-09-17		4	OVRLRESP	Overall Response	PR	2022-03-02
B2	2021-09-17		1	OVRLRESP	Overall Response	SD	2021-10-27
B2	2021-09-17		2	OVRLRESP	Overall Response	PR	2021-12-07
B2	2021-09-17		3	OVRLRESP	Overall Response	CR	2022-01-17
B3	2021-09-17		1	OVRLRESP	Overall Response	SD	2021-10-27
B3	2021-09-17		2	OVRLRESP	Overall Response	PR	2021-12-07
B3	2021-09-17		3	OVRLRESP	Overall Response	CR	2021-12-25
B3	2021-09-17		4	OVRLRESP	Overall Response	CR	2022-01-17
B4	2021-09-17		1	OVRLRESP	Overall Response	SD	2021-10-27
B4	2021-09-17		2	OVRLRESP	Overall Response	PR	2021-12-07
B4	2021-09-17		3	OVRLRESP	Overall Response	NE	2022-01-17
B4	2021-09-17		4	OVRLRESP	Overall Response	PR	2022-03-02
C1	2022-01-11		1	OVRLRESP	Overall Response	SD	2022-02-22
C1	2022-01-11		2	OVRLRESP	Overall Response	SD	2022-04-07
C1	2022-01-11		3	OVRLRESP	Overall Response	PD	2022-05-16
C2	2021-09-24		1	OVRLRESP	Overall Response	SD	2021-11-03
C2	2021-09-24		2	OVRLRESP	Overall Response	PR	2021-12-29
C2	2021-09-24		3	OVRLRESP	Overall Response	SD	2022-01-24
C2	2021-09-24		4	OVRLRESP	Overall Response	SD	2022-03-09
C3	2021-12-10		1	OVRLRESP	Overall Response	CR	2022-01-24
C3	2021-12-10		2	OVRLRESP	Overall Response	PD	2022-03-02
C4	2022-01-24		1	OVRLRESP	Overall Response	SD	2022-03-07
C5	2021-09-23		1	OVRLRESP	Overall Response	PR	2021-11-05
C5	2021-09-23		2	OVRLRESP	Overall Response	NE	2021-12-07
D1	2021-12-10		1	OVRLRESP	Overall Response	SD	2022-01-19
D1	2021-12-10		2	OVRLRESP	Overall Response	PD	2022-03-02
D2	2021-11-25		1	OVRLRESP	Overall Response	PD	2022-01-05
D2	2021-11-25		2	OVRLRESP	Overall Response	PR	2022-02-06
D2	2021-11-25		3	OVRLRESP	Overall Response	PR	2022-02-20
D3	2021-12-21	2022-02-12	1	OVRLRESP	Overall Response	PD	2022-01-20
D3	2021-12-21	2022-02-12	2	OVRLRESP	Overall Response	SD	2022-03-08
D4	2021-12-10		1	OVRLRESP	Overall Response	CR	2022-01-15
D4	2021-12-10		2	OVRLRESP	Overall Response	PD	2022-03-02
E1	2021-09-26		1	OVRLRESP	Overall Response	NE	2021-12-20
E2	2021-12-29		1	OVRLRESP	Overall Response	SD	2022-02-07
E3	2021-09-26		1	OVRLRESP	Overall Response	PR	2021-11-04
E3	2021-09-26		2	OVRLRESP	Overall Response	NE	2021-12-15

Table 2 WORK.RS0

Any assessments that are performed after the first documented PD or the start date of new anti-cancer therapy will not be considered in the BOR derivation. Example SAS codes are as below. Only assessments with ANL02FL="Y" would be used in the next data step.

```
data rs1;
    set rs0;
    format ADT yymmdd10.;
    ADT = input(RSDTC,yymmdd10.);
    if n(RANDDT,ADT)=2 then ADY =ADT-RANDDT+(ADT ge RANDDT);
run;

/* To select the 1st PD date from all assessments which are not after new
anti-cancer therapy (if any) */
proc sql;
    create table _1stPD as
    select distinct USUBJID,ADT as PDDT,RESEQ as PDSEQ
```

```
      from rs1 where RSSTRESC="PD" and ( . lt ADT le NCTXSDT or NCTXSDT=.)
      group by USUBJID
      having ADT=min(ADT)
      order by USUBJID;
   quit;

   /*ANL01FL -- to flag the first PD record*/
   data rs2;
      merge rs1(in=a) _1stPD;
      by USUBJID;
      if a;
      if RSSEQ=PDSEQ then ANL01FL="Y";
   run;

   /* ANL02FL – to flag the records that are not after first PD or new anti-
   cancer therapy start date (if any) */
   data rs3;
      set rs2;
      if not ( . lt NCTXSDT lt ADT ) and not ( . lt PDDT lt ADT ) then
   ANL02FL="Y";
   run;

   proc sort data=rs3;
     by USUBJID ADT;
   run;
```

After that, with using a RETAIN statement in the data step, the programming logics of CR/PR confirmation can be easily implemented. Example SAS codes, as well as the output data set Table 3 WORK.BOR1 are shown as below.

```
   /* CR/PR confirmation rules; SD minimum time interval from start date */
   data bor1;
      set rs3;
      by USUBJID ADT;
      where ANL02FL="Y";
      format ini_CRDT CRconfirmed CRDT ini_PRDT PRconfirmed PRDT SDDT;
      format ini_CRDT CRDT ini_PRDT PRDT SDDT yymmdd10.;
   /* CR */
      retain ini_CRDT;
      if first.USUBJID then ini_CRDT=.;
      if RSSTRESC="CR" and ini_CRDT=. then ini_CRDT=ADT;
      else if RSSTRESC not in ("CR","NE") then ini_CRDT=.;
      if RSSTRESC="CR" and ADT -ini_CRDT ge 28 and ini_CRDT ne . then do;
         CRconfirmed = 'Y';
         CRDT = ini_CRDT;
      end;
   /* PR */
      retain ini_PRDT;
      if first.USUBJID then ini_PRDT=.;
      if RSSTRESC="PR" and ini_PRDT=. then ini_PRDT=ADT;
      else if RSSTRESC not in ("CR","PR","NE") then ini_PRDT=.;
      if RSSTRESC in ("CR","PR") and ADT -ini_PRDT ge 28 and ini_PRDT ne .
   then do;
         PRconfirmed = 'Y';
         PRDT = ini_PRDT;
      end;
   /* SD */
```

```
        if RSSTRESC in ("CR","PR","SD") and (ADT-RANDDT ge 42) then SDDT=ADT;
        drop ini_CRDT ini_PRDT;
    run;
```

USUBJID	RANDDT	NCTXSDT	RSSTRESC	ADT	ADY	PDDT	ANL01FL	ANL02FL	CRconfirmed	CRDT	PRconfirmed	PRDT	SDDT
A1	2021-12-31		PR	2022-02-09	41			Y					
A1	2021-12-31		CR	2022-03-22	82			Y			Y	2022-02-09	2022-03-22
A1	2021-12-31		CR	2022-05-06	127			Y	Y	2022-03-22	Y	2022-02-09	2022-05-06
A2	2021-12-31		PR	2022-02-09	41			Y					
A2	2021-12-31		CR	2022-03-22	82			Y			Y	2022-02-09	2022-03-22
A2	2021-12-31		NE	2022-05-06	127			Y					
A2	2021-12-31		CR	2022-06-18	170			Y	Y	2022-03-22	Y	2022-02-09	2022-06-18
A3	2021-12-31		PR	2022-02-09	41			Y					
A3	2021-12-31		CR	2022-03-22	82			Y			Y	2022-02-09	2022-03-22
A3	2021-12-31		CR	2022-04-04	95			Y			Y	2022-02-09	2022-04-04
A3	2021-12-31		CR	2022-05-06	127			Y	Y	2022-03-22	Y	2022-02-09	2022-05-06
B1	2021-09-17		SD	2021-10-27	41			Y					
B1	2021-09-17		PR	2021-12-07	82			Y					2021-12-07
B1	2021-09-17		PR	2022-01-17	123			Y			Y	2021-12-07	2022-01-17
B1	2021-09-17		PR	2022-03-02	167			Y			Y	2021-12-07	2022-03-02
B2	2021-09-17		SD	2021-10-27	41			Y					
B2	2021-09-17		PR	2021-12-07	82			Y					2021-12-07
B2	2021-09-17		CR	2022-01-17	123			Y			Y	2021-12-07	2022-01-17
B3	2021-09-17		SD	2021-10-27	41			Y					
B3	2021-09-17		PR	2021-12-07	82			Y					2021-12-07
B3	2021-09-17		CR	2021-12-25	100			Y					2021-12-25
B3	2021-09-17		CR	2022-01-17	123			Y			Y	2021-12-07	2022-01-17
B4	2021-09-17		SD	2021-10-27	41			Y					
B4	2021-09-17		PR	2021-12-07	82			Y					2021-12-07
B4	2021-09-17		NE	2022-01-17	123			Y					
B4	2021-09-17		PR	2022-03-02	167			Y			Y	2021-12-07	2022-03-02
C1	2022-01-11		SD	2022-02-22	43	2022-05-16		Y					2022-02-22
C1	2022-01-11		SD	2022-04-07	87	2022-05-16		Y					2022-04-07
C1	2022-01-11		PD	2022-05-16	126	2022-05-16	Y	Y					
C2	2021-09-24		SD	2021-11-03	41			Y					
C2	2021-09-24		PR	2021-12-29	97			Y					2021-12-29
C2	2021-09-24		SD	2022-01-24	123			Y					2022-01-24
C2	2021-09-24		SD	2022-03-09	167			Y					2022-03-09
C3	2021-12-10		CR	2022-01-24	46	2022-03-02		Y					2022-01-24
C3	2021-12-10		PD	2022-03-02	83	2022-03-02	Y	Y					
C4	2022-01-24		SD	2022-03-07	43			Y					2022-03-07
C5	2021-09-23		PR	2021-11-05	44			Y					2021-11-05
C5	2021-09-23		NE	2021-12-07	76			Y					
D1	2021-12-10		SD	2022-01-19	41	2022-03-02		Y					
D1	2021-12-10		PD	2022-03-02	83	2022-03-02	Y	Y					
D2	2021-11-25		PD	2022-01-05	42	2022-01-05	Y	Y					
D3	2021-12-21	2022-02-12	PD	2022-01-20	31	2022-01-20	Y	Y					
D4	2021-12-10		CR	2022-01-15	37	2022-03-02		Y					
D4	2021-12-10		PD	2022-03-02	83	2022-03-02	Y	Y					
E1	2021-09-26		NE	2021-12-20	86			Y					
E2	2021-12-29		SD	2022-02-07	41			Y					
E3	2021-09-26		PR	2021-11-04	40			Y					
E3	2021-09-26		NE	2021-12-15	81			Y					

Table 3 WORK.BOR1

Based on WORK.BOR1, a data set containing CRDT, PRDT, SDDT, and PDDT is created. The data structure is one record per subject. Then, BOR for each patient can be determined in the order of CR>PR>SD>PD>NE. The example SAS codes together with the output data set Table 4 WORK.BOR3 are demonstrated as below.

```
proc sql;
    create table bor2 as
    select distinct a.USUBJID, a.FASFL
                ,min(b.CRDT) as CRDT format=is8601da.
                ,min(c.PRDT) as PRDT format=is8601da.
                ,min(c.SDDT) as SDDT format=is8601da.
                ,min(e.PDDT) as PDDT format=is8601da.
    from adsl(where=(FASFL='Y')) as a
    left join bor1 (where=(CRconfirmed = 'Y')) as b
        on a.USUBJID=b.USUBJID
    left join bor1 (where=(PRconfirmed = 'Y')) as c
        on a.USUBJID=c.USUBJID
    left join bor1 (where=(SDDT ne .)) as d
        on a.USUBJID=d.USUBJID
    left join bor1 (where=(PDDT ne .)) as e
        on a.USUBJID=e.USUBJID
    group by a.USUBJID
    ;
quit;
```

```
data bor3;
   set bor2;
   length PARAMCD $8  PARAM $200 AVALC $20;
   PARAMCD = "BESTRESP";
   PARAM = "Best Overall Response";
   if CRDT ne . then AVALC="CR";
   else if PRDT ne . then AVALC="PR";
   else if SDDT ne . then AVALC="SD";
   else if PDDT ne . then AVALC="PD";
   else AVALC="NE";
run;
```

USUBJID	FASFL	CRDT	PRDT	SDDT	PDDT	PARAMCD	PARAM	AVALC
A1	Y	2022-03-22	2022-02-09	2022-03-22		BESTRESP	Best Overall Response	CR
A2	Y	2022-03-22	2022-02-09	2022-03-22		BESTRESP	Best Overall Response	CR
A3	Y	2022-03-22	2022-02-09	2022-03-22		BESTRESP	Best Overall Response	CR
B1	Y		2021-12-07	2021-12-07		BESTRESP	Best Overall Response	PR
B2	Y		2021-12-07	2021-12-07		BESTRESP	Best Overall Response	PR
B3	Y		2021-12-07	2021-12-07		BESTRESP	Best Overall Response	PR
B4	Y		2021-12-07	2021-12-07		BESTRESP	Best Overall Response	PR
C1	Y			2022-02-22	2022-05-16	BESTRESP	Best Overall Response	SD
C2	Y			2021-12-29		BESTRESP	Best Overall Response	SD
C3	Y			2022-01-24	2022-03-02	BESTRESP	Best Overall Response	SD
C4	Y			2022-03-07		BESTRESP	Best Overall Response	SD
C5	Y			2021-11-05		BESTRESP	Best Overall Response	SD
D1	Y				2022-03-02	BESTRESP	Best Overall Response	PD
D2	Y				2022-01-05	BESTRESP	Best Overall Response	PD
D3	Y				2022-01-20	BESTRESP	Best Overall Response	PD
D4	Y				2022-03-02	BESTRESP	Best Overall Response	PD
E1	Y					BESTRESP	Best Overall Response	NE
E2	Y					BESTRESP	Best Overall Response	NE
E3	Y					BESTRESP	Best Overall Response	NE

Table 4 WORK.BOR3

Some sponsors/studies may pre-specified in the protocol/SAP that 1) the sequence of PR-SD-PR can be considered as a confirmed PR as long as the 2 PRs are observed at a minimum of 4 weeks apart, and/or 2) that the sequence of CR-PR-CR can be considered as a confirmed CR as long as the 2 CRs are observed at a minimum of 4 weeks apart and the intermediate PR is the case of target response being CR, non-target response being NE, and without any new lesions. In the example SAS codes of this paper, these 2 special cases are not taken into considerations.

HOW TO PREPARE ADAM DATA SETS FOR PFS ANALYSIS

Progression-free survival (PFS) is also commonly used as a primary or secondary endpoint in solid tumor clinical trials. PFS is defined as the time from randomization (or start of study treatment for non-randomized studies) until disease progression or death from any cause. The detailed PFS censoring schemes normally are specified in study statistical analysis plan (SAP). An example of PFS censoring schemes is shown as below Table 5, which comes from <Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics Guidance for Industry> released by FDA.

Situation	Date of Progression or Censoring	Outcome
Incomplete or no baseline tumor assessments	Randomization	Censored
Progression documented between scheduled visits	Earliest of: • Date of progression assessment showing new lesion (if progression is based on new lesion); or • Date of last progression assessment	Progressed
No progression	Date of last progression assessment with no documented progression	Censored
Treatment discontinuation for undocumented progression	Date of last progression assessment with no documented progression	Censored
Treatment discontinuation for toxicity or other reason	Date of last progression with no documented progression	*Censored*
New anticancer treatment started	Date of last progression assessment with documented nonprogression before start of new treatment	*Censored*
Death before first PD assessment	Date of death	Progressed
Death between adequate assessment visits	Date of death	Progressed
Death or progression after more than one missed visit	Date of last progression assessment with documented nonprogression	*Censored*

Table 5 Example of censoring schemes for PFS

Whereas, it's not easy for inexperienced statistical programmers to fully understand the reasons behind these schemes. The lack of particular basic knowledge in oncology therapeutic area will puzzle programmers. Questions could be like:

1. Patients without any event will be censored at the last date when progression status is adequately assessed instead of censored at the cutoff or dropout date. Why?

2. When documented progression or death happens after patient missing scheduled assessments, patients will be classified as censoring instead of event. Why?

3. There are so many dates to be considered, including progression date, death date, last adequately assessed date, and new anti-cancer therapy start date. If two or more dates happen to one patient, situation becomes complicated. How to deal with this?

4. How to translate the PFS schemes into programming logic to ensure they are mutually exclusive and completely exhaustive?

To understand the PFS censoring schemes, firstly, we should know that, for patients with solid tumors, post-baseline overall responses are determined via radiologic scanning on tumors, which normally are performed by fixed-interval schedules as per the protocol, e.g., every 6 weeks. Thus, progression assessments are not observed continuously. Progression actually occurs between two radiologic assessments and we are not able to know the exact moment when it occurred. In patients with no evidence of progression before data cutoff or dropout, we are unable to know the exact tumor status at the date of data cutoff or dropout date. Therefore, such patients will be censored at the last date on which progression status was adequately assessed.

For some patients, progression or death may occur after missed scheduled tumor radiological assessments. Progression was possibly already detected earlier if previous radiological assessments not missed. In this case, to avoid potential bias, patient will be regarded as censored at the date of last adequate tumor assessment, instead of an event.

New anti-cancer therapy may confound the relationship between the trial drug and PFS outcome. Thus, radiological assessments after the new anti-cancer therapy start date will not be considered in the PFS algorithms. Patient will be censored at the date of the last adequate tumor assessment that explicitly documented no progression prior to the start of the new anti-cancer therapy.

For demonstration, mockup ADSL, ADRS and ADDATES data sets are created.

Example ADSL data set is shown in below Table 6.

USUBJID	FASFL	BLTUASFL	RANDDT	NCTXSDT	DTHDT
Unique Subject Identifier	Full Analysis Set Population Flag	Baseline Tumor Assessments Flag	Date of Randomization	Start Date of New Anti-cancer Therapy	Date of Death
A	Y	N	2021-01-01		
B	Y	Y	2021-02-01		
C	Y	Y	2021-03-01		
D	Y	Y	2021-04-01	2021-09-01	
E	Y	Y	2021-05-01	2021-06-01	
F	Y	Y	2021-06-01		2021-12-01
G	Y	Y	2021-07-01		2021-10-01
H	Y	Y	2021-07-02		
I	Y	Y	2021-07-03		

Table 6 ADSL

Example ADRS data set is shown in below Table 7. In the ADRS, any assessments that are performed after the first documented PD or the start date of new anti-cancer therapy will not be considered in the PFS schemes. ANL01FL is used to flag the first PD record and ANL02FL to flag all assessments that are not after the first documented PD or the start date of new anti-cancer therapy. For ANL01FL and ANL02FL, programming logics are the same with the above example for BOR derivation.

USUBJID	ASEQ	PARAMCD	PARAM	VISIT	AVALC	ADT	ANL01FL	ANL02FL
Unique Subject Identifier	Analysis Sequence Number	Parameter Code	Parameter	Visit	Analysis Value (C)	Analysis Date	Analysis Flag 01	Analysis Flag 02
B	1	OVRLRESP	Overall Response	Visit 1	SD	2021-03-12		Y
B	2	OVRLRESP	Overall Response	Visit 2	NE	2021-04-24		Y
B	3	OVRLRESP	Overall Response	Visit 3	SD	2021-06-05		Y
B	4	OVRLRESP	Overall Response	Visit 4	SD	2021-07-18		Y
B	5	OVRLRESP	Overall Response	Visit 5	PD	2021-10-30	Y	Y
B	5	OVRLRESP	Overall Response	Visit 6	PD	2021-11-30		
C	1	OVRLRESP	Overall Response	Visit 1	PD	2021-04-13	Y	Y
D	1	OVRLRESP	Overall Response	Visit 1	SD	2021-05-11		Y
D	2	OVRLRESP	Overall Response	Visit 2	SD	2021-06-23		Y
D	3	OVRLRESP	Overall Response	Visit 3	SD	2021-08-07		Y
D	4	OVRLRESP	Overall Response	Visit 4	SD	2021-09-19		
F	1	OVRLRESP	Overall Response	Visit 1	SD	2021-07-12		Y
F	2	OVRLRESP	Overall Response	Visit 2	SD	2021-08-23		Y
F	3	OVRLRESP	Overall Response	Visit 3	SD	2021-10-05		Y
F	3	OVRLRESP	Overall Response	Visit 4	SD	2021-11-18		Y
G	1	OVRLRESP	Overall Response	Visit 1	NE	2021-08-13		Y
H	1	OVRLRESP	Overall Response	Visit 1	PR	2021-08-13		Y
H	2	OVRLRESP	Overall Response	Visit 2	PR	2021-09-25		Y
H	3	OVRLRESP	Overall Response	Visit 3	SD	2021-11-10		Y
H	4	OVRLRESP	Overall Response	Visit 4	SD	2022-12-20		Y

Table 7 ADRS

Example ADDATES, an intermediate data set, is shown in below Table 8. RANDDT, NCTXSDT and DTHDT come from ADSL. PDDT is ADRS.ADT where ANL01FL="Y". LSASSDT is the maximum ADT from ADRS records where ANL02FL="Y" and AVALC not in ("NE", "PD").

USUBJID	ASEQ	ADT	ADTDESCD	ADTDESC
Unique Subject Identifier	Analysis Sequence Number	Analysis Date	Description of Analysis Date Code	Description of Analysis Date
A	1	2021-01-01	RANDDT	Date of Randomization
B	1	2021-02-01	RANDDT	Date of Randomization
B	2	2021-07-18	LSASSDT	Date of Last Adequate Tumor Assessment
B	3	2021-10-30	PDDT	Date of Documented Progression
C	1	2021-03-01	RANDDT	Date of Randomization
C	2	2021-04-13	PDDT	Date of Documented Progression
D	1	2021-04-01	RANDDT	Date of Randomization
D	2	2021-08-07	LSASSDT	Date of Last Adequate Tumor Assessment
D	3	2021-09-01	NCTXSDT	Start Date of New Anti-cancer Therapy
E	1	2021-05-01	RANDDT	Date of Randomization
E	2	2021-06-01	NCTXSDT	Start Date of New Anti-cancer Therapy
F	1	2021-06-01	RANDDT	Date of Randomization
F	2	2021-11-18	LSASSDT	Date of Last Adequate Tumor Assessment
F	3	2021-12-01	DTHDT	Date of Death
G	1	2021-07-01	RANDDT	Date of Randomization
G	2	2021-10-01	DTHDT	Date of Death
H	1	2021-07-02	RANDDT	Date of Randomization
H	2	2021-12-20	LSASSDT	Date of Last Adequate Tumor Assessment
I	1	2021-07-03	RANDDT	Date of Randomization

Table 8 ADDATES

After ADDATES is transposed, each important date is represented by a variable, as shown in below Table 9.

USUBJID	FASFL	BLTUASFL	RANDDT	LSASSDT	PDDT	NCTXSDT	DTHDT
Unique Subject Identifier	Full Analysis Set Population Flag	Baseline Tumor Assessments Flag	Date of Randomization	Date of Last Adequate Tumor Assessment	Date of Documented Progression	Start Date of New Anti-cancer Therapy	Date of Death
A	Y	N	2021-01-01				
B	Y	Y	2021-02-01	2021-07-18	2021-10-30		
C	Y	Y	2021-03-01		2021-04-13		
D	Y	Y	2021-04-01	2021-08-07		2021-09-01	
E	Y	Y	2021-05-01			2021-06-01	
F	Y	Y	2021-06-01	2021-11-18			2021-12-01
G	Y	Y	2021-07-01				2021-10-01
H	Y	Y	2021-07-02	2021-12-20			
I	Y	Y	2021-07-03				

Table 9 WORK.DATES

According to the example PFS censoring schemes, patients can be classified into 9 different types which are mutually exclusive and completely exhaustive, as shown in below Figure 1.
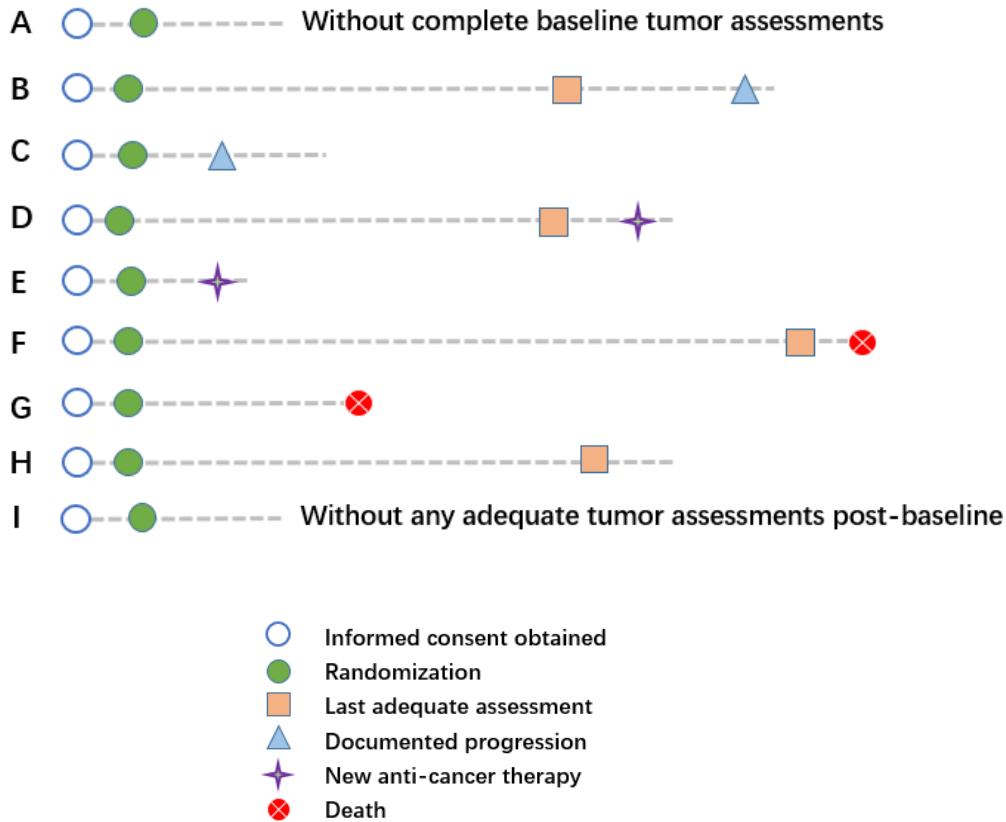
A ○--● ------------ Without complete baseline tumor assessments

B ○-● ------------------------- ▭ ----------- ▲

C ○-● ------- ▲ -----

D ○-● ------------------- ▭ --- ✦

E ○-● ---- ✦

F ○-● --------------------------------- ▭ --❌

G ○-● ------- ❌

H ○-● ------------------ ▭ ----

I ○--● -------- Without any adequate tumor assessments post-baseline

○ Informed consent obtained
● Randomization
▭ Last adequate assessment
▲ Documented progression
✦ New anti-cancer therapy
❌ Death

Figure 1 9 different cases of PFS censoring schemes

Depending on whether a patient had missed more than 1 scheduled visit (84 days is used in the example, normally the maximum time interval will be pre-specified in the SAP), patients of type B/C/F/G are classified into 2 sub-types. Example SAS codes are provided as below.

```
data adtte;
   set work.dates;
   length PARAMCD $8 PARAM $200 EVNTDESC CNSDTDSC $200;
   format STARTDT ADT YYMMDD10.;
   PARAMCD = "PFS";
   PARAM= "Progression Free Survival (months)";
   STARTDT = RANDDT;
   if FASFL="Y";
/* Type A -- Incomplete or no baseline tumor assessments */
   if BLTUASFL ='N' then do;
      CNSR = 1; ADT = RANDDT; EVNTDESC="Incomplete or no baseline tumor
assessments";
      CNSDTDSC="Date of Randomization";
   end;
/* Type B1 -- PD between scheduled visits, and there is at least 1 adequate
tumor assessments before PD */
   else if PDDT ne . and . lt PDDT-LSASSDT le 84 then do;
      CNSR = 0; ADT = PDDT; EVNTDESC="Documented Progression";
   end;
/* Type B2 -- PD after more than one missed visit, and there is at least 1
adequate tumor assessments before PD */
   else if PDDT ne . and PDDT-LSASSDT gt 84 then do;
```

```
      CNSR = 1; ADT = LSASSDT; EVNTDESC="Progression after more than one
missed visit";
      CNSDTDSC="Date of last assessment with documented nonprogression";
   end;
/* Type C1 -- PD between scheduled visits, and there is no adequate tumor
assessments before PD */
   else if PDDT ne . and LSASSDT=. and . lt PDDT-RANDDT le 84 then do;
      CNSR = 0; ADT = PDDT; EVNTDESC="Documented Progression";
   end;
/* Type C2 -- PD after more than 1 missed visit, and there is no adequate
tumor assessments before PD */
   else if PDDT ne . and LSASSDT=. and PDDT-RANDDT gt 84 then do;
      CNSR = 1; ADT = LSASSDT; EVNTDESC="Progression after more than one
missed visit";
      CNSDTDSC="Date of last assessment with documented nonprogression";
   end;
/* Type D -- New anti-cancer therapy started, and there is at least 1
adequate tumor assessment before that*/
   else if PDDT= . and NCTXSDT ne . and LSASSDT ne . then do;
      CNSR = 1; ADT = LSASSDT; EVNTDESC="New anti-cancer therapy started";
      CNSDTDSC="Date of last assessment with documented nonprogression
before start of new treatment";
   end;
/* Type E -- New anti-cancer therapy started, and there is no adequate
tumor assessment before that */
   else if PDDT= . and NCTXSDT ne . and LSASSDT=. then do;
      CNSR = 1; ADT = RANDDT; EVNTDESC="New anti-cancer therapy started";
      CNSDTDSC="Date of Randomization";
   end;
/* Type F1 -- Dead between scheduled visits, and there is at least 1
adequate tumor assessment before death */
   else if PDDT= . and NCTXSDT=. and DTHDT ne . and LSASSDT ne . and . lt
DTHDT-LSASSDT le 84 then do;
      CNSR = 0; ADT = DTHDT; EVNTDESC="Death";
   end;
/* Type F2 -- Dead after more than 1 missed visit, and there is at least 1
adequate tumor assessment before death */
   else if PDDT= . and NCTXSDT=. and DTHDT ne . and LSASSDT ne . and DTHDT-
LSASSDT gt 84 then do;
      CNSR = 1; ADT = LSASSDT; EVNTDESC="Death after more than one missed
visit";
      CNSDTDSC="Date of Last Adequate Tumor Assessment";
   end;
/* Type G1 -- Dead between scheduled visits, and there is no adequate tumor
assessment before death */
   else if PDDT= . and NCTXSDT=. and DTHDT ne . and LSASSDT=. and . lt
DTHDT-RANDDT le 84 then do;
      CNSR = 0; ADT = DTHDT; EVNTDESC="Death";
   end;
/* Type G2 -- Dead after more than one missed visit, and there is no
adequate tumor assessment before death */
   else if PDDT= . and NCTXSDT=. and DTHDT ne . and LSASSDT=. and DTHDT-
RANDDT gt 84 then do;
      CNSR = 1; ADT = RANDDT; EVNTDESC="Death after more than one missed
visit";
      CNSDTDSC="Date of Randomization";
   end;
```

```
/* Type H -- no PD/death/new anti-cancer therapy, and there is at least 1
adequate tumor assessment post-baseline */
    else if PDDT= . and NCTXSDT=. and DTHDT= . and LSASSDT ne . then do;
        CNSR = 1; ADT = LSASSDT; EVNTDESC="No Progression";
        CNSDTDSC="Date of Last Adequate Tumor Assessment";
    end;
/* Type I -- no PD/death/new anti-cancer therapy, and there is no adequate
tumor assessment post-baseline*/
    else if PDDT= . and NCTXSDT=. and DTHDT= . and LSASSDT= . then do;
        CNSR = 1; ADT = RANDDT; EVNTDESC="Without post-baseline tumor
assessments";
        CNSDTDSC="Date of Randomization";
    end;
    AVAL = (ADT-STARTDT+1)/30.4375;
run;
```

CONCLUSION

In this paper, basic knowledge on BOR and PFS are explained. Additionally, with using examples, the algorithms of how to derive confirmed BOR and PFS are also demonstrated.

Readers should notice that all the example data sets for demonstrations are intermediate data sets instead of final ADaM datasets conform to CDISC. Example SAS programs cannot be directly used before adaptions have been made according to a specific study protocol and SAP.

REFERENCES

Eisenhauer EA, Therasse P, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009 Jan;45(2):228-47. doi: 10.1016/j.ejca.2008.10.026. PMID: 19097774.

Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics Guidance for Industry, FDA, 2015, available at https://www.fda.gov/media/116860/download.

CONTACT INFORMATION

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

<Sen Fan>
<CSPC Pharmaceutical Group Ltd>
<+86 18810488457>
< fansen_vincent@163.com >