PhUSE US Connect 2019

Paper DH13

Compliance Report for Patient-Reported Outcomes

Amy Chu, Merck Research Laboratories, Merck & Co., Inc., Rahway, NJ USA
Christine Teng, Merck Research Laboratories, Merck & Co., Inc., Rahway, NJ USA
Jia He, Merck Research Laboratories, Merck & Co., Inc., Rahway, NJ USA

ABSTRACT

Patient-Reported Outcomes (PRO) are becoming an important component in the evaluation of prescription drugs in registration, reimbursement and labeling claims. A PRO instrument (i.e., a questionnaire plus the documentation that support its use) is a means to capture data used to measure treatment benefit or risk in medical product for clinical trials. It is an important outcome measure that in addition to the classical outcomes, such as tumor control and survival, cancer trials should also evaluate the effect of a treatment it has on a patient's quality of life.

One of the major concerns about PRO analysis is missing data. Missing data leads to decrease in precision and power caused by the reduction in data, and the potential for bias in analysis. This paper will focus on the summary of compliance/completion report, and the ADaM dataset design to capture the missing reasons in the report.

SAS®9, Windows, Intermediate Level
Key Words: Patient-Reported Outcomes, PRO, QS, questionnaire, ADaM

INTRODUCTION

PRO endpoints in the field of oncology are neither pure efficacy nor pure safety endpoints as they are affected by both disease progression and treatment tolerability. PRO analysis is getting more attention in recent years for drug approval. PRO analysis in oncology can be challenging due to missing data. There are various reasons for PRO missing data. Some missing reasons were collected/pre-defined, and other examples of missing data occur when subjects fail to report visits or visits are discontinued before the end of the clinical trial.

We will go through the potential reasons for missing PRO data, corresponding ADaM data structure design, and the sample report which displays the missing reasons to show compliance and completion rate by time point. The compliance report is important to determine the time point when the analysis of PRO data can be analyzed with meaningful outcome.

REASONS FOR MISSING DATA

There are many reasons for missing PRO data. In most cases, the vendor will obtain missing reasons as pre-defined. However, there are other cases of missing data which may be due to reasons such as: subject discontinued treatment, subject died, or vendor data not transferred. To evaluate the missing PRO reasons correctly, SDTM domains SV (Subjects visits) and DS (Disposition) data would be considered. Below are the six possible reason categories for missing modes including the pre-defined rules. Appendix section has samples of data display for different scenarios and the flow chart when to derive the reasons.

1) **Vendor data**:

Site will collect missing mode on subject visits. Some examples of the pre-specified reasons for collected missing mode reasons include:

- 'NOT COMPLETED DUE TO SITE STAFF ERROR'
- 'SUBJECT IN HOSPITAL OR HOSPICE'
- 'TRANSLATION NOT AVAILABLE IN SUBJECTS LANGUAGE'
- 'SUBJECT REFUSED FOR OTHER REASONS'
- 'SUBJECT WAS PHYSICALLY UNABLE TO COMPLETE'

2) **End of treatment**:
In cases where site did not collect missing mode, treatment discontinuation reason is considered if available. For example, Subject discontinued the study for reasons such as subject withdrew consent or clinical progression. This case subject is not able to continue the required visits and unable to completed PRO instruments. Some examples of treatment discontinuation reasons to use for missing reason on a given timepoint include:

- 'ADVERSE EVENT'
- 'CLINICAL PROGRESSION
- 'WITHDRAW BY SUBJECTS'
- 'DEATH'

The information is collected in the SDTM DS domain (subject disposition). See case #1 under appendix1.

3) **Subject has died:**
At the time of death, the information was not yet entered in the disposition status nor instruments.

4) **With visit, no record**:
Subject attended the visit but missing PRO data for that visit due to vendor data has not been transferred or entered. See case #2 under appendix 1.

5) **Visit not scheduled**:
In the defined windows, there is no scheduled visit nor PRO data. See case #3 under appendix 1.

6) **Visit not reached:**
Subject has not yet reached the defined visit window. See case #2 under appendix 1.

ADaM DESIGN

The CDISC data structure used to create the PRO analysis is Basic Data Structure (BDS). BDS is an analysis dataset that contains one or more records per subject, per analysis parameter, per analysis time point. The parameters are mapped from –TEST and –TESTCD, and additional parameters are created as needed for derived scores and in our case, also the compliance missing reasons.

The following approach is intent to follow CDISC data structure, data variable requirement and to provide traceability from the analysis dataset back to their corresponding records in the SDTM domain.

Dataset Variable and Structure Conventions

In the QS domain, PRO status and missing reasons data records should be carried to the analysis dataset, if records exist. QSTEST, QSTESTCD should be copied directly into PARAM and PARAMCD. QSSTRESC containing PRO status and missing reasons should also be copied to the analysis dataset variable AVALC and AVAL. In addition to the original PRO compliance status records, separate PRO status records should then be created as the derived compliance status in each defined analysis window. We suggest using PARAMTYP with value 'DERIVED' for the derived PRO compliance records and assign unique PARAM and PARAMCD values, so that any derived PRO compliance records can be easily traced back to their original QS status records or end of treatment data, if it exists. For example, the records based on PRO vendor collected missing reason as QSTESTCD ='QUESTAT', PARAMCD value would be QUESTAT. The derived PRO status record with PARAMTYP='DERIVED' and PARAMCD='QUECOMP' would be created for the analysis dataset and used for the compliance analysis, detail

sample will be provided in the appendix section of the paper. For maximum traceability, SRC*XXX* variables such as SRCDOM, SRCVAR and SRCSEQ can be utilized in the analysis dataset. In studies where more than one PRO instrument is used, categorization variable PARCATy can be used for the purpose. Below table is a partial example for the proposed key variables and example derivations:

Variable Name	Variable Label	Type	Define Derivation
PARAM	Parameter	Char	QS.QSTEST For derived records, please refer to analysis plan for detail logics.
PARAMCD	Parameter Code	Char	QS.QSTESTCD For derived records, please refer to analysis plan for detail logics.
PARCAT1	Parameter Category 1	Char	QS.QSCAT
PARAMTYP	Parameter Type	Char	DERIVED' for derived compliance records
AVISIT	Analysis Visit	Char	Avisit represents the analysis visit for the record. Using ADY, low-1="BASELINE" 2-10="WEEK 3" 11-20="WEEK 6" 21-30="WEEK 9" 31-high="OutofFlow";
ADY	Analysis Relative Day	Num	Example derivation: ADT-ADSL.TRTSDT + 1 if ADT>= TRTSDT, else ADT-ADSL.TRTSDT if ADT < TRTSDT
AVALC	Analysis Value (C)	Char	QS.QSSTRESC For PARAMTYP ='DERIVED' compliance records, please refer to analysis plan for detail logics.
AVAL	Analysis Value	Num	Numeric representation of AVALC
AVALCAT1	Analysis Category 1	Char	Example Derivation: if AVAL = 1 then AVALCAT1 = 'Missing by design' if AVAL = 2 then AVALCAT1 = 'Expected to complete Questionnaires'
SRCDOM	Source Domain	Char	Example Derivation: For PARAMTYP ='DERIVED', if QS missing reason is used then SRCDOM='QS'; if DS end of treatment reason is used then SRCDOM='DS';
SRCVAR	Source Variable	Char	Example Derivation: For PARAMTYP ='DERIVED', if QS missing reason is used then SRCVAR='QSSTRESC'; if DS end of treatment reason is used then SRCVAR='DSDECOD';
SRCSEQ	Source Sequence Number	Num	Example Derivation: For PARAMTYP ='DERIVED', if QS missing reason is used then SRCSEQ='QSSEQ'; if DS end of treatment reason is used then SRCSEQ='DSSEQ';

Use of AVISIT with SAS Format

In the PRO analysis dataset, analysis visits AVISIT variable are often required. To ensure the analysis window derivation are flexible and consistent across the study, we recommend using format for the derivation. The format approach allows the pre-defined window range to be easily applied and updated across study. This practice also ensures the windows are applied consistently across the study. In cases where there are multiple records within the same window, analysis flag ANLzzFL with value Y can be used to identify the records to be used in analysis.

Similarly, we also recommend that SAS format be used for AVAL and AVALC for missing PRO reasons. Each reason is assigned with a unique AVAL and matching AVALC in the PRO missing reason format. This practice allows for

quicker modification when new reason is added or updated, also allows easier modification when table format need to be adjusted.

Below is a sample code to create the format, to be utilize with the combination of input or put function in data steps, such as avisit=put(ady, avisit.) or avisitn=input(ady, avisitn), similar logic can be applied to AVAL and AVALC;

```
proc format;
       invalue avisitn
       low-1   =1
       2-10    =2
       11-20   =3
       21-30   =4
       31-high=999;

       value avisit
       low-1="BASELINE"
       2-10="WEEK 3"
       11-20="WEEK 6"
       21-30="WEEK 9"
       31-high="OutofFlow";

run;
```

Use of AVALCAT1

Another important aspect of designing the analysis dataset is how the analysis dataset can be mapped into proper categorical variables for the analysis tables. Most often PRO missing reasons are divided in different categories for the analysis. In our approach, we suggest the use of AVALCAT1 for the categorical representation of AVAL/AVALC. Using AVALCAT1, we can map AVALC into two main categories,

1) **Missing by Design**, reasons include but not limited to subject discontinued treatment, subject did not reach or scheduled the visits, or subject has died.
2) **Expected to complete Questionnaires**, include reasons when subject is unable to fill out the form for another reason or vendor data has not been transferred or obtained. This category also included subjects completed the questionnaires.

COMPLETEON/COMPLIANCE RATE

Completion Rate is defined as the percentage of the number of subjects who complete at least one item over the number of subjects in the population at each analysis window

$$Completion\ Rate = \frac{Number\ of\ Subjects\ who\ Complete\ at\ least\ one\ Item}{Number\ of\ Subjects\ in\ the\ population}$$

The completion rate is expected to shrink in the later visit during study period due to the subjects who discontinued early or data issues. Therefore, another measurement, Compliance Rate, defined as the percentage of observed visits over number of eligible subjects who are expected to complete the PRO assessment.

$$Compliance\ Rate = \frac{Number\ of\ Subjects\ who\ Complete\ at\ least\ one\ Item}{Number\ of\ Eligible\ Subjects\ who\ are\ Expected\ to\ Complete}$$

Missing data can cause low compliance/completion rate and potentially lead to bias. This makes the analysis difficult to assess the effect between treatments and subject's quality of life. Moreover, the lower compliance/completion rate would indicate the PRO endpoint as unreliable. To effectively confirm the reliability of PRO analysis, the proportion at the visit with overall completion/compliance rates are suggested to reach ~60%/~80%, respectively.

COMPLIANCE REPORT EXAMPLE

The example below is a sample of the EQ-5D compliance table. Since we have all visit and compliance derived in analysis dataset, the table counts are direct frequency counts. The PRO compliance tables are counts by treatment visit with the reasons mentioned above. Two categories are displayed, missing by design and expected to complete questionnaires. Based on the table, we can calculate the compliance and completion rate by visit to assess the effect and reliability of PRO analysis. Please note the table below is a mock-up for illustration with sample reasons to demonstrate the difference in completion rate vs. compliance rate. All reasons should be included in final report even has 0 counts.

Compliance of EQ-5D-3L by Visit and by Treatment
(FAS Population)

Treatment Visit	Category	Treatment A N = 50		Treatment B N = 42	
		n	(%)	n	(%)
Week 6	Missing by design	8	(16.0)	29	(69.1)
	Discontinued due to adverse event	2	(4.0)	16	(38.1)
	Discontinued due to clinical progression	2	(4.0)	11	(26.2)
	Discontinued due to death	0	(0.0)	0	(0.0)
	Discontinued due to withdraw by subject	0	(0.0)	0	(0.0)
	Translation not available in subjects language	2	(4.0)	1	(2.4)
	Visit not scheduled	2	(4.0)	1	(2.4)
	Visit not reached	0	(0.0)	0	(0.0)
	Subject died, no reported disposition status	0	(0.0)	0	(0.0)
	Expected to complete questionnaires	42	(84.0)	13	(30.9)
	Not completed	2	(4.0)	2	(4.8)
	Subject in hospital or hospice	1	(2.0)	0	(0.0)
	Subject refused for other reasons	1	(2.0)	2	(4.8)
	Not completed due to site staff error	0	(0.0)	0	(0.0)
	Subject was physically unable to complete	0	(0.0)	0	(0.0)
	With visit no record	0	(0.0)	0	(0.0)
	Completed	40	(80.0)	11	(26.2)
	Compliance (completed per protocol)*	40	(95.2)	11	(84.6)

* Compliance is the proportion of subjects who completed the PRO questionnaire among these who are expected to complete at each time point, excluding these missing by design.

OTHER CONSIDERATIONS

Whether to include questionnaire response and compliance in the same analysis dataset, or separate into different analysis datasets, should be decided by the company standard or study team. Defining each miss mode can be tricky, it is important to discuss and decide the logic based on the study PRO analysis design and data collection to ensure report accuracy and reduce bias. All questionnaire missing reason mode and analysis window rules should be documented in the define.xml.

SUMMARY

In summary, to provide a meaningful PRO analysis, it is important to understand the reasons of missing data. We want to minimize missing data due to study design, administration or the complexity of questionnaire. At the beginning of study, non-avoidable missing data should be evaluated, and pre-specified rules should be developed to catch the missing reasons. In this paper, we lay out examples of some missing reasons and how we utilize CDISC ADaM BDS structure and variables to create the compliance report to monitor the missing data.

REFFERENCES

Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims – FDA December 2009
https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf

ACKNOWLEGEMENTS

The author would like to thank their management team for encouragement and review of this paper.

CONTACT INFORMATION

Amy Chu Merck Research Laboratories, Merck & Co., Inc. shu-ching.chu@merck.com	
Christine Teng Merck Research Laboratories, Merck & Co., Inc. christine_teng@merck.com	sas® Certified Advanced Programmer
Jia He Merck Research Laboratories, Merck & Co., Inc. jia.he@merck.com	

APPENDIX1: ADaM DESIGN DATA STRUCTURE EXAMPLES

Case 1: The subject discontinued the study with in week 18 (visit window) due to progressive disease, hereafter, the PRO missing reason is discontinued due to clinical progress.

USUBJID	PARAMCD	PARAMTYP	AVISIT	AVALC	AVALCAT1	SRCDOM	SRCVAR
1111	QUECOMP	DERIVED	BASELINE	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
1111	QUECOMP	DERIVED	WEEK 3	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
1111	QUECOMP	DERIVED	WEEK 6	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
1111	QUECOMP	DERIVED	WEEK 9	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
1111	QUECOMP	DERIVED	WEEK 18	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 27	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 36	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 45	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 54	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD

1111	QUECOMP	DERIVED	WEEK 66	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 78	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 90	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD
1111	QUECOMP	DERIVED	WEEK 102	DISCONTINUED DUE TO CLINICAL PROGRESSION	Missing by Design	DS	DSDECOD

Case 2: The subject with visit at week 45 in SV domain but no corresponding PRO data in the defined visit window, this case "With visit, no record" will be used as missing reason. Considering week 45 is last visit for this subject, the following visits are not reached (Missing by Design).

USUBJID	PARAMCD	PARAMTYP	AVISIT	AVALC	AVALCAT1	SRCDOM	SRCVAR
2222	QUECOMP	DERIVED	WEEK 36	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
2222	QUECOMP	DERIVED	WEEK 45	WITH VISIT, NO RECORD	Expected to Complete Questionnaires		
2222	QUECOMP	DERIVED	WEEK 54	VISIT NOT REACHED	Missing by Design		
2222	QUECOMP	DERIVED	WEEK 66	VISIT NOT REACHED	Missing by Design		
2222	QUECOMP	DERIVED	WEEK 78	VISIT NOT REACHED	Missing by Design		
2222	QUECOMP	DERIVED	WEEK 90	VISIT NOT REACHED	Missing by Design		
2222	QUECOMP	DERIVED	WEEK 102	VISIT NOT REACHED	Missing by Design		

Case 3: Visit not scheduled can be easily confused with visit not reached, the difference can be distinguished by identifying whether the visit is subject's latest visit or between other visits. If between visits with the condition that the subject has no collected PRO data and did not attend the required visit within the defined window, then the reason will be given as "Visit not scheduled" as week 3 below.

USUBJID	PARAMCD	PARAMTYP	AVISIT	AVALC	AVALCAT1	SRCDOM	SRCVAR
3333	QUECOMP	DERIVED	BASELINE	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC
3333	QUECOMP	DERIVED	WEEK 3	VISIT NOT SCHEDULED	Missing by Design		
3333	QUECOMP	DERIVED	WEEK 6	COMPLETED	Expected to Complete Questionnaires	QS	QSSTRESC

APPENDIX2: FLOWCHART FOR MISSING REASON DERIVATION