PharmaSUG China 2023 - Paper DS-135

Data Handling of Rollover Subjects in Master Protocol Design

Fu Wang, Mona Mo, Parexel International

ABSTRACT

This paper provides an overview and implementation details of data handling for rollover subjects in oncology studies with master protocol design. The basic idea of master protocol study design and data structure of related CRF/SDTM/ADaM/TFL are discussed and attention points provided for related coding activities.

INTRODUCTION

Expediting drug development and approval process to better serve patients has been a challenge for the industry and regulators around the world alike. Developing trial designs that can test multiple drugs and/or multiple cancer subgroups parallelly under a single protocol without a need to develop new protocols for each trial is one of the approaches that could expediate drug development process and has attracted increased attention within clinical trial industry in recent years. The term master protocol is often used to describe innovative design of such trials, with a variety of terms such as umbrella, basket or platform describing specific designs.

In general, a master protocol is defined as a protocol designed with multiple sub studies, which may have different objectives and involve coordinated efforts to evaluate one or more investigational drugs in one or more disease subtypes within the overall trial structure[1]. Umbrella trial is designed to study multiple targeted therapies in the context of a single disease, basket trial is designed to study single targeted therapy for multiple diseases or disease subtypes, and platform trial is designed to study multiple target therapies for a single disease in a perpetual manner with therapies allowed to enter and leave platform based on certain decision algorithm[2].

SCENARIO

During clinical trials, it's possible that subjects may roll over from one study to another within the same product, and there is a chance that subjects may roll over from one subprotocol to another subprotocol within the same master protocol. Furthermore, subjects may roll over more than once, for example from one study to another study for the same product followed by a second rollover from one subprotocol to another. Unlike screen failure data which in most cases would not be included in analyses, rollover subjects are just like any subjects enrolled in each study or each subprotocol, would accumulate large volume of data in safety and/or efficacy to support analyses.

With the advance of clinical research and sound scientific justification, subject rollover from study to study or from subprotocol to subprotocol is not uncommon, which poses a challenge to statistical programming in terms of data handling and compliance of industrial standards. Specifically, how should we present our data in USUBJID/SUBJID and related datasets for subjects with rollover history which not only can meet the study analysis needs but also in compliance with regulatory agency requirements and industry standards as much as possible?

CDISC NOTES AND US FDA Technical Conformance Guide (TCG)

USUBJID and SUBJID are two subject identifier variables in CDISC SDTM domains and ADaM datasets. In short, the USUBJID is a unique identifier in product/submission level while SUBJID is a unique identifier in study/trial level and the original notes[3] from CDISC v3.2 provided the following details.

USUBJID: *Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. This must be a unique number, and could be a compound identifier formed by concatenating STUDYID-SITEID-SUBJID.*

SUBJID: *Subject identifier, which must be unique within the study. Often the ID of the subject as recorded on a CRF.*

However, CDISC didn't provide specific guidelines on how to handle subjects with rollover situations, but FDA TCG does provide detailed instructions[4] in section 4.1.1.2 as below which could be followed as a good practice.

SUBJID: *If a single subject is screened and/or enrolled more than once in a study, then the subject's SUBJID should be different for each unique screening or enrollment. For a study with multiple screenings and/or multiple enrollments per subject, SUBJID should be included in other related domains besides DM even though it may cause validation errors.*

USUBJID: *Each individual subject should be assigned a single unique identifier across the entire application. An individual subject should have the exact same unique identifier across all datasets, including between SDTM and ADaM datasets. Subjects that participate in more than one study should maintain the same USUBJID across all studies.*

IMPLEMENTATION

Based on the information from CDISC and FDA TCG as mentioned above, we would like to illustrate how we handled rollover subjects in one of our projects as follows.

CRF Design

To properly handle data of rollover subjects, the first thing was how to define the value of USUBJID. Our study team decided to use and retain the value of USUBJID for the subject when the subject was first successfully enrolled in the study of the product. And to keep traceability between different studies or within the same master protocol study, we designed a form outlined below in **Table 1** to collect previous study/subprotocol information as needed.

(*Assuming only study Z is a master protocol design here*)

Fields	Options
Prior Study Number	X|Y|Z
Prior Subprotocol	A|B|C (for study Z only)
Prior Subject ID	Free text

Table 1 Rollover History

SDTM/ADaM Data Structure

In dataset level, we have both variables for USUBJID and SUBJID not only in SDTM.DM following CDISC standards, but also other related domains as well per previous mentioned FDA TCG. Example of data structure is provided below, provided that SUBJID is consist of SITEID-NUM and USUBJID consists of STUDYID-SITEID-NUM.

For subjects just rollover once, we have data structure showed below in Table 2.

Row 1, rollover between studies, the subject was first enrolled in study X with USUBJID as X-01-001 and then rolled over to study Z subprotocol A with SUBJID as 02-002, and we also can trace back to study X with USUBJID X-01-001 in product level submission package.

Row 2, rollover between studies, like row 1, the subject was first enrolled in study Y with USUBJID as Y-03-003 and then rolled over to study Z subprotocol B with SUBJID as 04-004.

Row 3 and 4, rollover between subprotocols within master protocol study, the subject was first enrolled in study Z subprotocol A with USUBJID as Z-05-005 and then rolled over to another subprotocol C within study Z. Note that for row 3, the USUBJID only contains information of one more part from STUDYID and

the rest two parts are the same with SUBJID in the form of SITEID-NUM (05-005) and for row 3 & 4, the USUBJID is kept the same while only the SUBJID was changed from 05-005 to 08-008.

ROW	STUDYID	USUBJID	SUBJID	SUBPRT
1	Z	X-01-001	02-002	A
2	Z	Y-03-003	04-004	B
3	Z	Z-05-005	05-005	A
4	Z	Z-05-005	08-008	C

Table 2 Data structure in study Z for subjects who rollover once

For subjects who rollover twice, we have data structure showed below in Table 3.

Row 1 and 2, rollover between studies and second rollover between subprotocols within master protocol study, the subject was first enrolled in study X with USUBJID as X-07-007 and then rolled over to study Z subprotocol A with SUBJID as 10-010, after that, the subject then rolled over to subprotocol B within study Z with SUBJID as 12-012, and we also can trace back to study X with USUBJID X-07-007 in product level submission package.

Row 3 and 4, rollover between studies and second rollover between subprotocols within master protocol study, like above example, the subject was first enrolled in study Y with USUBJID as Y-09-009 and then rolled over to study Z subprotocol A with SUBJID as 14-014 followed by a second rollover to subprotocol C within study Z with new SUBJID as 16-016.

ROW	STUDYID	USUBJID	SUBJID	SUBPRT
1	Z	X-07-007	10-010	A
2	Z	X-07-007	12-012	B
3	Z	Y-09-009	14-014	A
4	Z	Y-09-009	16-016	C

Table 3 Data structure in study Z for subjects who rollover twice

As we can see from above examples that, current approach is adding variable of SUBJID into other related SDTM domains besides SDTM.DM which is in alignment with current FDA TCG although a minor deviation from CDISC standards so far, and corresponding P21E validation warning messages can be properly explained as needed.

TFL Outputs

When it comes to the outputs for analyses, we can handle it properly based on different requests we received.

For a single subprotocol analysis, we can filter corresponding subprotocol subjects in advance and no impact caused by rollover subjects at all.

For combined subprotocol analysis, such as immuno oncology combination, if the outputs are needed only in subprotocol level and no need to analyze data across subprotocols, then still no impact from rollover subjects.

For integrated summary of safety and/or efficacy analyses, then we need to consult with study team and/or the regulatory agency whether we want to report in SUBJID level or USUBJID level since for rollover subjects, their SUBJID are different within same USUBJID.

Programming and P21E Considerations

From the perspective of master protocol study, when rollover happens with subjects, USUBJID itself seems no longer unique within the master protocol study level, since SUBJID could be varied for each successful enrollment, causing that for a single individual subject, the USUBJID can correspond to different SUBJID in data. However, from the point of product level or entire submission/application perspective, we are still having USUBJID as the unique identifier which can be used for better traceability.

During the coding process, our recommendation is to always consider USUBJID and SUBJID together when performing DATA MERGE steps and/or PROC SQL joint steps. Moreover, the logic of EPOCH/SEQ/BLFL variables should also be considered with these two together.

For P21E validation, we will get a warning message as 'Variable appears in dataset but is not in SDTM model' stated in SD0058, and another message as 'Duplicate USUBJID in DM within STUDYID' stated in SD0083 which both can be properly explained per FDA guidelines.

CONCLUSION

Master protocol design has gained increased interest in biopharmaceutical industry and has started changing the landscape of clinical research. With master protocol design, handling rollover subjects poses a unique challenge in statistical programming. The approach outlined in the paper provides an example on how to tackle this challenge and is fully compliant with US FDA Technical Conformance Guide, is easy to implement in programming, and meets the various analysis requirements.

REFERENCES

1 "Master Protocols: Efficient Clinical Trial Design Strategies to Expedite Development of Oncology Drugs and Biologics Guidance for Industry", Accessed 13May2023 https://www.fda.gov/regulatory-information/search-fda-guidance-documents/master-protocols-efficient-clinical-trial-design-strategies-expedite-development-oncology-drugs-and

2 Woodcock J, LaVange LM. "Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both." N Engl J Med. 2017 Jul 6;377(1):62-70. doi: 10.1056/NEJMra1510062. PMID: 28679092.

3 CDISC. "SDTMIG v3.2" https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v3-2

4 FDA "Study Data Technical Conformance Guide", Accessed 13May2023 https://www.fda.gov/regulatory-information/search-fda-guidance-documents/study-data-technical-conformance-guide-technical-specifications-document

CONTACT INFORMATION

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

Fu Wang
Parexel International
+86-176-2190-3558
Fu.Wang@parexel.com

Mona Mo
Parexel International
Mona.Mo@parexel.com