

Considerations When Representing Multiple Subject Enrollments in SDTM

Kristin Kelly, Pinnacle 21, LLC

PhUSE US Connect 2019

Paper DS14

Agenda

Introduction

Background

What Do Sponsors Do?

What Might Work Better?

So.....What's Next?

Conclusion

Introduction (1)

- When a subject is permitted to re-screen or re-enroll in a clinical trial or program, there are few scenarios that can occur depending on the study protocol (listed most to least common):
 - 1) Subjects that enroll in more than one study.
 - 2) Subjects that are screened more than once within the same study.
 - 3) Subjects that enroll more than once and are treated within the same study.
- Other scenarios that can happen that are not outlined in the protocol:
 - Subjects that change sites during a trial
 - Subjects that enroll at different sites unbeknownst to the investigator
- Regardless of whether the re-enrollment is planned or unplanned, it is at the discretion of the sponsor how to collect enough data that will support past instances of a subject in a study
 - e.g. Informed consent dates, previous STUDYID and SUBJID, etc.

Introduction (2)

- Currently, the SDTMIG does not have guidance or standards in place to handle this scenario of multiple subject instances.
- Because of this, sponsors have chosen different ways to represent this data where some are problematic because:
 - The data is not SDTM-conformant
 - Does not load into regulatory review tools
 - Is not in keeping with current regulatory guidance
- These issues with the submitted data lengthen the review time needed.
- In this presentation, the following will be discussed:
 - Examples from industry that may not work so well
 - An interim solution for imminent submissions that conform with the SDTM standard as it is today

Background (1)

- Per the FDA Technical Conformance Guide (TCG), each subject should be assigned the same unique identifier (USUBJID) across the entire submission even if they participate in more than one study
- All subjects should have a SUBJID assigned that is unique within the study
 - Also stated in the TCG: ‘if a subject is screened more than once in a trial, then the subject’s SUBJID should be different’ for each screening attempt
- There is typically a one-to-one relationship between USUBJID and SUBJID for subjects that appear once in a submission
 - This is not the case when collecting data for re-screened subjects or subjects that enroll more than once in the same study or across studies

Background (2)

- Even in the case of multiple instances of a subject across the submission, the same USUBJID value should be maintained but the SUBJID values should be different.
- To better understand this concept, it may be helpful to think of the same USUBJID being assigned to the same person or ‘warm body’ regardless of how many times they appear in a submission.
- For example, each person will still have the same characteristics each time they are screened or enrolled,
 - i.e. the same race, ethnicity, sex, birthdate etc.
- SUBJID is simply a number assigned to that same person that is different each time they screen for a clinical trial.

Background (3)

- Further, the TCG states in Section 4.1.1.3, *‘in the DM domain, each subject should have only one single record per study.’*
- DM Domain, Assumption #2 from SDTMIG v3.2:
 - *‘If a subject changes sites (or participates more than once) in the same study, ‘the sponsor must decide how to populate variables such as USUBJID, SUBJID, and SITEID based on their operational and analysis needs, but only one DM record should be submitted for the subject.’*
- Combining these concepts when dealing with multiple enrollments can become difficult when it comes time to prepare submission data that is:
 - Transparent
 - SDTM-conformant
 - Can easily load into review tools.

What Do Sponsors Do?

- It is a common misconception that most study designs do not allow for multiple enrollments, especially within the same study.
 - Occurs often enough that sponsors have asked for submission guidance on how to model this data.
- In March 2014, there was a PhUSE WG that created solutions for modeling this data but it would require an update to the SDTM standard
 - The recommendations were the following:
 - 1) Design the CRF to handle collection of data for previous screenings/enrollments
 - 2) Create more than one record in DM for each screening/enrollment event per subject
 - a) The subject would have the same USUBJID but the SUBJID values would be different
 - 3) Add SUBJID to the general observation class domains as a permissible Identifier in order to differentiate records specific to each instance in the trial

Scenario #1: Subjects That Re-screen Or Re-enroll Within The Same Study (1)

- Example study, XYZ1047, is a Phase III, global trial for the investigational vaccine, Kryptonex
- Allows subjects to re-screen multiple times until they are enrolled
- Subject 505 at Site #2 is assigned SUBJID = '002-0505' but fails screening
- The subject returns for another screening attempt and is assigned SUBJID = '002-0618' but fails screening a second time



Scenario #1: Subjects That Re-screen Or Re-enroll Within The Same Study (2)

- The subject returns for a third attempt and is assigned SUBJID = '002-0725'
- This time the subject meets all criteria and is randomized and subsequently treated
- Assigned a USUBJID that reflects the SUBJID value from the last screening attempt
 - USUBJID = 'XYZ1047-002-0725'



Scenario #1: Subjects That Re-screen Or Re-enroll Within The Same Study (3)

- Based on the recommendations mentioned, the DM domain contained a record for each screening attempt.

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC		RFICDTC	SITEID	ARMCD
XYZ1047	DM	XYZ1047-002-0725	002-0505				2017-04-23	002	SCRNFAIL
XYZ1047	DM	XYZ1047-002-0725	002-0618				2017-04-30	002	SCRNFAIL
XYZ1047	DM	XYZ1047-002-0725	002-0725	2017-05-23	2018-10-23		2017-05-14	002	XYZ1047

- Same USUBJID assigned for all records
- SUBJID values are different for each attempt
- RFSTDTC/RFENDTC is only populated for the instance where the subject was randomized and treated
- SUBJID was added as a supplemental qualifier to general observation class domains rather than the parent domain.

Scenario #1: Subjects That Re-screen Or Re-enroll Within The Same Study (4)

- Applying this approach in DM triggers the following validation rule:

FDA Validator Rule ID	FDA Validator Message	Publisher	Publisher ID	Business or Conformance Rule validated	Validator Rule
SD0083	Duplicate USUBJID	CDISC	CG0151	Identifier used to uniquely identify a subject across all studies	The value of Unique Subject Identifier (USUBJID) variable must be unique for each subject across all trials in the submission.

- The sponsor provided an explanation in the cSDRG:
 - ‘This study had subjects that re-screened more than once until they were eligible. Each screening attempt is represented as a record in DM. This is according to the recommendation of the USUBJID PhuSE WG that models multiple records in DM.’

*Please note: SD0083 is looking for more than one record in DM within one study. It does not currently check across studies.

Scenario #1: Subjects That Re-screen Or Re-enroll Within The Same Study (5)

- Everything seems in place until.....
 - The study is received at the regulatory agency and the data does not load into review tools

How can this be? The sponsor explained what they did in the cSDRG!

- The tools were expecting one record per subject in DM
- Please note: this does not make the data unreviewable, but....
 - It does increase the review time because alternative approaches need to be used to review and load the data
- Bottom line (literally): Unless the sponsor has checked with their review division: **Don't Do This!**

Scenario #2: Subjects That Enroll in More than One Study (1)

- Much more common when encountering multiple subject instances
- As stated, USUBJID should be unique for the same person across a submission
- At times, some sponsors have difficulty either:
 - Tracking previous enrollments
 - Lack the ability to assign a unique identifier due to tool restrictions
- These reasons may or may not be mutually exclusive.
 - Examples include:
 - Studies submitted were conducted before CDISC standards were required and the data was not collected
 - A study drug purchased from another sponsor and not all the information needed to track these subjects is available.

Scenario #2: Subjects That Enroll in More than One Study (2)

- In this example, the sponsor did not collect data from previous enrollments and was unable to assign a unique USUBJID:
- Subject 202, enrolled and completed Study ABC709:

Study ABC709 – *dm.xpt*

STUDYID	USUBJID	SUBJID	RFSTDTC	RFENDTC		BRTHDTC	SEX	RACE	ARMCD
ABC709	ABC709-08-0202	08-0202	2016-10-08	2017-04-25		1955	F	WHITE	ABC_DRUG

- Then enrolled and completed Study ABC710:

Study ABC710 – *dm.xpt*

STUDYID	USUBJID	SUBJID	RFSTDTC	RFENDTC		BRTHDTC	SEX	RACE	ARMCD
ABC710	ABC710-14-0560	14-0560	2017-10-14	2018-08-23		1955	F	WHITE	ABC_DRUG

Scenario #2: Subjects That Enroll in More than One Study (3)

- Because this subject was not assigned the same USUBJID, the reviewer did not readily know that each study enrollment was the same person. The only way to determine this would be:
 - To compare the subject's characteristics such as BIRTHDATE, SEX, RACE, ETHNIC etc.
 - Subject's height and weight may also be needed
- Not being able to track a subject throughout a submission could be problematic for safety and efficacy analyses, as well as determination of long-term effects
- It is important to employ:
 - Good data collection practice for previous enrollments
 - A process for assigning the same USUBJID to the same person

What Should Sponsors Do (That Might Work Better)?

- Ideally, the SDTM standard would have been updated to incorporate the proposed recommendations discussed earlier by allowing multiple records in DM or some other proposed strategy that makes sense.
- Standards development and regulatory acceptance move much slower
- What can be done NOW that will still achieve conformance with current guidance?

Scenario #1 (Revisited) – Multiple Enrollments in the Same Study (1)

- Recall this example of a subject that failed screening twice before being eligible at the third screening attempt.
- To be conformant with the SDTMIG and the TCG, there should be one record per subject:

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC		RFICDTC	SITEID	ARMCD
XYZ1047	DM	XYZ1047-002-0725	002-0725	2017-05-23	2018-10-23		2017-05-14	002	XYZ1047

- Previous SUBJID values can be stored in SUPPDM:

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
XYZ1047	DM	XYZ1047-002-0725			PSUBJID	Previous Subject Identifier	002-0505
XYZ1047	DM	XYZ1047-002-0725			PSUBJID	Previous Subject Identifier	002-0618

- Subject characteristics in DM such as RACE, ETHNIC, SEX, BRTHDTC, etc that will not change for each screening do not need to be duplicated in SUPPDM

Scenario #1 (Revisited) – Multiple Enrollments in the Same Study (2)

- Previous informed consent dates as well as the primary informed consent when the subject enrolled can be stored in the DS domain:

STUDYID	DOMAIN	USUBJID	DSSEQ	DSTERM	DSDECOD	DSCAT	DSSTDTC
XYZ1047	DS	XYZ1047-002-0725	1	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	2017-04-23
XYZ1047	DS	XYZ1047-002-0725	2	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	2017-04-30
XYZ1047	DS	XYZ1047-002-0725	3	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	2017-05-14

- In order to differentiate between each event, the corresponding SUBJID can be stored in SUPPDS.

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
XYZ1047	DS	XYZ1047-002-0725	DSSEQ	1	SUBJID	Subject Identifier	002-0505
XYZ1047	DS	XYZ1047-002-0725	DSSEQ	2	SUBJID	Subject Identifier	002-0618
XYZ1047	DS	XYZ1047-002-0725	DSSEQ	3	SUBJID	Subject Identifier	002-0725

- Same approach can be used for other subject-level domains.

Scenario #1 (Revisited) – Multiple Enrollments in the Same Study (3)

- Though this approach may not be viewed as the ‘ideal’ it does provide the following:
 - SDTM conformance
 - Conformance to the Technical Conformance Guide
 - Data that readily loads into regulatory review tools
- Some ‘not so ideal’ points to consider for this approach:
 - If there are many previous dates stored in SUPPDM, this data is not machine-readable as a supplemental qualifier, e.g. ‘Informed Consent Date 1’, ‘Informed Consent Date 2’, etc.
 - Recommend storing these dates elsewhere perhaps in the domain in which they should reside such as DS
 - Does not work well when a subject is treated and completes the study more than once (that may occur in trials for rare diseases or low subject availability)
 - Because only one record is in DM, it forces the sponsor to choose the ‘primary’ enrollment when each time a subject is treated and passes through the trial is important

Scenario #2 (Revisited): Subjects That Enroll in More than One Study (1)

- Recall this example, the sponsor did not collect data from previous enrollments and was unable to assign a unique USUBJID.
- Ideally, the sponsor would have planned to collect this data and would be able to take the time before the submission to re-assign USUBJID for subjects that are determined to be the same person

Study ABC709 – *dm.xpt*

STUDYID	USUBJID	SUBJID	RFSTDTC	RFENDTC		BRTHDTC	SEX	RACE	ARMCD
ABC709	ABC709-08-0202	08-0202	2016-10-08	2017-04-25		1955	F	WHITE	ABC_DRUG

Study ABC710 – *dm.xpt*

STUDYID	USUBJID	SUBJID	RFSTDTC	RFENDTC		BRTHDTC	SEX	RACE	ARMCD
ABC710	ABC709-08-0202	14-0560	2017-10-14	2018-08-23		1955	F	WHITE	ABC_DRUG

So.....What's Next? - A Final Thought on the TCG (1)

- In the most recent version of the TCG, there was language added on how to handle multiple enrollments that would still mandate having one record per subject in Demographics (DM):
 - *‘For subjects with multiple enrollments within a single study, the primary enrollment should be submitted in DM. Additional enrollments should be included in a custom domain with a similar structure to DM. Clarifying statements in the RG would be helpful.*
 - *For subjects with multiple screenings and no subsequent enrollment, include the primary screening in DM with additional screenings in a custom domain with a structure similar to DM.*
 - *For subjects with multiple screenings and subsequent enrollment, include the enrollment in DM with screenings in a custom domain with a structure similar to DM.’*

So.....What's Next? - A Final Thought on the TCG (2)

- It is requested that for multiple enrollment data that a custom domain for a domain classified as Special Purpose be created.
 - Per the SDTMIG, custom domains can only be created based on one of the three general observation classes: Events, Interventions, or Findings.
 - Following this request would require modeling data that is not in keeping with the SDTMIG.
 - Also, it is unclear what variables from DM should be included in this custom domain.
- It is also requested that sponsors choose a 'primary' enrollment or screening (if the subject is not enrolled)
 - It is not readily apparent how a primary record could be determined
- Until the requirements are provided and the standard is updated, it is recommended to apply the approach discussed here that would ensure conformance with the SDTM standard as well as communicate with the review division, if necessary.

Conclusion

- Allowing multiple subject instances either within the same study or across studies has become quite common
- Currently, the SDTM standard nor regulatory agencies provide guidance on how to submit this data, but there have been a few proposals from the FDA and a PhUSE Working Group.
- Though these approaches seem straightforward, they may not fully comply with the standard as it exists today.
- Submitting data that is not in compliance with SDTM may increase review times if the data cannot readily load into regulatory review tools.
- Until the SDTM standard is updated, it may be helpful for sponsors to adopt an approach that is in keeping with the standard in order to shorten review times and to get drugs to patients faster.

Questions?

Name: Kristin Kelly

Organization: Pinnacle 21, LLC

E-mail: kkelly@pinnacle21.com