Answers to Remaining Questions

Webinar: CDISC Members-Only Mini-Training: Ensuring USUBJID is Unique for an Individual in an Application

Presenters: Monica Mattson, Celgene; Mark Sullivan, PAREXEL; Gary Walker, Quintiles

Date: 12 Feb 2015

1. The Technical Conformance Guide recommendation for multiple USUBJID records in DM conflicts with FDA validation rule FDAC041 - so we can't submit this way yet, correct?

Monica Mattson: First, the Technical Conformance Guide states that a separate SUBJID should be assigned for every screening event. It is logical that there would be multiple USUBJID records in DM, though not specifically stated in the Technical Conformance Guide. Allowing multiple records per USUBJID in the DM domain is the recommendation of the Working Group. Second, I'm not familiar with FDAC041, but assume it checks for one record per USUBJID in the DM domain. A submission that uses multiple records per USUBJID in the DM domain would indeed fail that validation rule. If a sponsor wanted to send a submission that contains multiple records per USUBJID in the DM domain, they would need to discuss and get approval from their Review Division prior to doing so.

2. Regarding using the first USUBJID, there might be an issue in the subsequent study, if that USUBJID is already assigned to another patient in the subsequent study. In such cases, should we be using the last USUBJID?

Gary Walker: I think that using the first screening attempt SUBJID to populate the USUBJID is smart, if you concatenate STUDYID || SITEID || SUBJID to populate USUBJID. There is no requirement to concatenate these variables to create USUBJID, and each sponsor is left to their own devices to create their method of populating USUBJID. The key to populating USUBJID is to be sure that the value is unique to a person, whether the person participates in multiple trials or participates multiple times in one trial.

3. Are there any considerations on how multiple records per subject impacts ADSL which also has a 1 record per subject requirement?

Monica Mattson: The potential effect of the Working Group's recommendations on any other CDISC model besides SDTM was specifically determined to be out-of-scope for the purposes of our evaluation.

4. What is your recommendation for identifying rescreened subjects under the current SDTM guidelines?
Mark Sullivan: There is nothing in the current SDTM guidelines that deals with rescreened subjects. I would expect you could follow the Working Group’s recommendation to identify a subject’s previous SITEID and SUBJID so as to establish USUBJID.

Monica Mattson: I agree with Mark. I would also add that, most commonly, it would be expected that subjects would rescreen with the same investigator, and the investigator would be aware of the subject’s previous SITEID and SUBJID.

5. What does the subject participation database capture in order to identify unique humans, gender, initials, birthdate, age, height, weight, etc.?

Mark Sullivan: The USUBJID Working Group has not made any specific recommendations for a subject participation database, but certainly gender, initials, birthdate, age, height, weight or any demographic or physical characteristics would go a long way toward identifying subjects who have participated in the same study more than once. A combination of the CRF data and some programmed checks of basic demographic and physical attributes is that basic due diligence that is incumbent on the sponsor. However, the Working Group leaves it to each sponsor to determine what kind of tracking system they want to put in place and what processes they need to put in place.

6. Mark: With the recommendation to have multiple records in DM domain per USUBJID, is it your recommendation to just provide any explanation to OpenCDISC warnings/errors in a study data reviewer’s guide?

Mark Sullivan: At this time, the recommendation to allow multiple records per USUBJID in DM is under review and consideration. The only recommendation that should be freely implemented now, if you so choose, is to include a CRF where you ask if the subject participated in this study or a study in the same application.

7. Can a flag be created in SUPPQUAL for the domains that collected previous screen data, so it won’t break the SDTM standard for DM domain as one record per USUBJID?

Gary Walker: A flag in SUPPQUAL could be used to identify which records belong to screen failures, but some studies allow subjects to complete the trial and then re-enroll (for example, ). A flag in SUPPQUAL would be inadequate to address this rather complex data presentation problem.

8. It may be difficult to remain compliant with 21 CFR Part 11 to maintain data authenticity and traceability to source records, if USUBJID data are stored at the site, in query/edit check/EDC/IVR... systems and at Lab/ECG vendors. How can you synchronize this change universally in paper health records and within all of these systems when data are already historically stored?

Mark Sullivan: USUBJID is only meant to identify an individual person across all studies within an application. Hence the only place one would expect to see USUBJID is in the submitted datasets. There is no compelling reason for USUBJID to be in source records and, if it were, there is no reason it needs to be synchronized to the application’s datasets.
9. The same person is randomized twice so he/she can have two USUBJID in DM?

Gary Walker: Per SDTM rules, no person can have more than one USUBJID value. The USUBJID was designed to represent a person and the value is unique to the person.

10. Do you recommend to use DM to collect SUBJID from different trials? If not, what domain will this CRF be modeled to?

Gary Walker: CDASH does talk to the collection of prior trial identifiers in the DM domain. CDASH included USUBJID as a reminder to collect sufficient information about other trial participation to properly populate USUBJID. This recommendation in CDASH is in the DM domain section.

11. How realistic is it to expect subjects to know what their subject number was in a previous study?

Mark Sullivan: It is difficult to say, since it is dependent on what is shared between the site and the subject. I believe it is unlikely that a subject will know this information, but either the subject or the site could request that information from the previous site.

12. Why does the recommendation from the working violate current rules? I am not sure that sponsors would want to go that route. Are you that opposed to leaving prior SUBJID information in SUPPDM?

Monica Mattson: The Working Group's recommendation is not aligned with the common practice of defining the key structure of the DM domain ("one record per subject") as one record per USUBJID. The Working Group initially considered leaving prior SUBJID information in the SUPPDM dataset. However, when we created examples of how the datasets would merge, we found the result to be unwieldy. Implementation of this recommendation would necessarily re-designing the rules to reflect the change.

13. We have begun searching for suspected unique individuals by identifying subjects across multiple studies who have the same demographic profile (birthdate, sex, race, ethnicity) who have different USUBJIDs. We then ask the sponsor if they might be the same person. Could this be a useful strategy before data submission in identifying unique subjects?

Mark Sullivan: This could be a very useful strategy to find study subjects that are potentially the same individual. The Working group discussed this as well, but chose not to include this in the White Paper, because each sponsor will choose to implement this differently.

Monica Mattson: An additional consideration is that, with HIPAA and the implementation of patient privacy laws, the regular availability of detailed demographic information such as birthdate is less common than in the past.

14. Would a subject know his previous trial number and SUBJID? Even if he says he was in a previous trial, how could you identify the previous SUBJID when we can't collect a full DOB?

Mark Sullivan: We acknowledge that it's unlikely that a subject would know a previous SITEID and/or SUBJID. I believe it is incumbent upon the subject to contact the previous site and
request the information, or identify the previous investigator so that the current site can follow up either directly or through a study monitor to try to determine the previous SITEID and SUBJID.

15. How can data of different SUBJIDs (belonging to the same USUBJID) in SUPPDM be distinguished? (we have been using QNAM = SUBJID)

Gary Walker: Putting SUBJID in SUPPDM is problematic, as shown in the webinar's slides, because the SUBJID value is just one value and there are other, associated values (such as RFICDTC) that should be related to each SUBJID value. If a person has more than two SUBJID values, there is no way to correlate the proper SUBJID with its own variable values like RFICDTC (see slides for detailed example).

16. Should IC dates be in DS and not DM?

Gary Walker: Since SDTM-IG v3.1.2 amendment 1, informed consent is a variable in DM (RFICDTC) and should be in DM. There might also be a record in DS for the protocol milestone indicating when the informed consent was signed, but having that record in DS is a sponsor decision. Putting the informed consent date in DM (in RFICDTC) is not a sponsor decision, as RFICDTC is a required field.

17. Can multiple records in DM with same USUBJID lead to open CDISC errors or warnings?

Monica Mattson: Currently, multiple records in DM with the same USUBJID will lead to an OpenCDISC error. Implementation of the Working Group's recommendation would necessitate re-designing the rules to reflect the change.

18. Given HIPPA & identity laws, is it likely a sponsor will be able to collect & store patient information to identify patients across studies? Aside from the CRF displayed, more data would be needed.

Mark Sullivan: The assumption is that the site is not collecting more information for this purpose. At very least, the answers to the questions on the CRF would cause some additional work on the part of the sponsor to try to determine what SUBJIDs should be associated with the same USUBJID. Ideally, if they can be identified, it will make the linking easier. Further, if it were not evident, then some demographic and physical characteristics could be used to limit the number of possible matches. Working with site monitors may help identify the links. Remember that what we said was that some due diligence is expected on the sponsor's part, but we recognize that this may not catch every instance of USUBJID linked to an individual.

19. Was consideration given to how multiple records for a single subject who completed two different trials would be presented in an integrated SDTM database?

Monica Mattson: The handling of USUBJID in an integrated SDTM database was considered out-of-scope for the Working Group. However, it did come up during discussions, and the preliminary thoughts were (1) assuming that SUBJID is unique in a study, then the variables STUDYID and SUBJID could be concatenated to identify a unique screening event in an integrated database; (2) USUBJID could be used to identify an individual person.
20. When do you propose talking to the regulators? Which mechanism should be used and at what point in time?

Monica Mattson: This topic was considered to be out-of-scope for the USUBJID Working Group. However, generally speaking, the recommendation frequently presented, in multiple forums, is to speak with the Review Division as early as possible regarding standards-related questions.

21. How necessary do the panel members feel that screen fail attempts and data be included in submissions?

Monica Mattson: Whether screen failure data should be included in an application should be discussed with your Review Division. If there is agreement not to include screen failure data in the application, then the assignment of USUBJID is anticipated to be more straightforward, based on collection of the information needed to determine whether an individual has enrolled in any other study within the application. Data collection issues are addressed in the first recommendation of the Working Group.

22. Did you consider whether FA could be used to hold all the earlier screening information? That might be a way to hold all those different dates. Though I agree the solution you chose seems to work quite nicely, I didn't see that this was an option you considered.

Gary Walker: This approach was eliminated early in the Working Group's deliberations, as the Findings About (FA) domain is designed to hold data that represent findings about an Event or Intervention. FA was never intended to hold findings about DM (or any other special purpose domain).

23. In case due diligence fail, can demographic information (sex, DOB, initials) be used to identify and re-assign USUBJID at submission time?

Monica Mattson: The objective of the USUBJID Working Group was to find the best solution to ensure a unique USUBJID value for an individual with data in an application. We felt this would be best accomplished by identifying USUBJID as early as possible during a study (as outlined in our first recommendation). Additionally, with HIPAA and the implementation of patient privacy laws, the regular availability of detailed demographic information such as DOB and initials is less common than in the past.

24. Why couldn't FDA just have a recommendation that sponsors implement a 'unique human' identifier in the integrated datasets for ISE and ISS? For the CDISC group, their recommended variable name and discussion would come as part of recommendations for 'integrated' data?

Monica Mattson: This discussion was not in scope for the USUBJID Working Group. However, I would venture to guess that, while this solution may work in some cases, it is not a universal solution as it does not address studies that are not integrated as part of an application.

25. Why not create a Screening domain? This is an atypical situation so problems like the VS data shown with an odd IC date would have to be understood, or explained in a Review Guide.

Gary Walker: The team started with the premise that the current SDTM/SDTM-IG, with the current domains, were appropriate for the process we were designing. The problem is not that
there are data that don't fit in currently modeled domains, but that the data are "atypical" from the perspective of rescreening. There might be good reason for the FDA reviewers to look at the screening failure data (or for combining data for the same person in the same study where the subject who completes is allowed to enroll again as a new subject) and having the data in the standard domains actually makes the review easier than if the data were in other domains.