Approaches to Creating ADaM Subject-Level Analysis Datasets (ADSL) for Integrated Analyses

CDISC NJ Users Group
2014-01-14
Nate Freimark
Theorem Clinical Research
The SDTM and ADaM standards currently cover single study data structures in great detail. What has not yet been covered explicitly by the CDISC standards is how to combine data for an integrated analysis. Whether the target is an Integrated Summary of Safety (ISS) or an Integrated Summary of Efficacy (ISE), a great deal of planning is required to create a set of data that will support the desired analyses. While industry and the CDISC teams are looking into some answers and guidance around integrated data, the current need for integrated data has been pointed to by the FDA review divisions.
• In addition to ongoing considerations within the SDTM and ADaM teams, integration was the focus of one of the working groups that was identified as being needed at the 2012 FDA/PhUSE Computational Science Symposium that occurred March 19th and 20th, 2012.

• A Working Group was formed to address “Challenges of Integrating and Converting Data across Studies”.

• Until an industry consensus is developed, there are many decisions that must be made regarding the path that will be taken to get to the integrated analyses required.

• Consensus has not yet occurred
ADSL - Not Without Challenges

• Where to Begin
• Source Data
• One ADSL or Two
• Number of Records per Subject
• Update or Create New Variables
• Age, Treatment and Other Recalculated Variables
Which End Is Up?

• **Start at the beginning**
  - Gather what was done
  - Do I have study-level SDTM/ADaM?
  - Will they support my integrated analyses?
  - Determine if ADaM recreation is desired on a study level
  - Build on top of each step as needed
  - Good luck integrating everything 😊

• **Start at the end**
  - Determine ISS/ISE analysis needs
  - Make sure study level SDTM/ADaM are properly populated for integrated analyses
What is the Basis?

• **Source Data**
  - Study SDTM
  - Integrated SDTM
  - Study ADaM
Study Level SDTM

• **Pros**
  - Source data for ADSL remains unchanged and matches the source data for the individual studies
  - No additional SDTM programming or decisions are needed

• **Cons**
  - Events that cross studies are not combined
  - VISIT/TPT/TESTCD and other values are not consistent
  - Traceability can be difficult
  - Recoding would need to be done at a study level, for example, bringing each study MedDRA coding up to a more recent version used in the later studies
Integrated SDTM

• **Pros**
  - Source data is updated to support ADaM values
  - Define.xml can point to data within same study instead of looking back at multiple versions of the same domain across the individual studies for the ADaM value source
  - Allows for capturing in SDTM data that may not be available in a single record or domain in the individual studies
    - Medical history for extension studies
    - AE end dates for roll-over studies
    - Demographics not collected in roll-over studies
• **Cons**
  - Requires an extra level of dataset creation and QC
  - Requires harmonization of SDTM variables, values, and structures
    - VISIT/VISITNUM
    - TPT/TPTNUM
    - TEST/TESTCD
Study Level ADaM

**Pros**
- All calculated values are available
- Structures are similar
- Allows for availability of ADaM-derived values on an integrated level
  - Study level population flags
  - Disallowed medication flags (Did subject take a prohibited medication (captured in study ADCM))
- Exposure duration/compliance values

**Cons**
- Requires 100% pre-planning
- Requires addition of variables on a study level that are not relevant on a study-level analysis
- Assumes that all integrated ADaM will have a study-level ADaM
How Many ADSL?

• **Multiple**
  - ISS
    - Typically includes all studies with subgroups focused on Disposition, Exposure and AE groupings
  - ISE
    - Typically a focused set of studies with many more variables and subgroups

• **Single**
  - ISS/ISE
    - Are the same rules used for both?
    - How to handle populations that do not exist in all studies?
    - Dosing date vs. randomization date as anchors
How Many Records Per Subject?

- ADaM model default expectation is one record per subject
  - Based on single study design
- CDER Common Issues Document - “Integrated summaries may contain more than one record per unique subject in the case that an individual subject was enrolled in more than one study.”
How Many Records Per Subject? (cont)

- **One Record Per Subject**
  - Redefine TRTSDT/TRTEDT if a subject is in more than one study
  - Much traceability gets lost in the transition

- **One Record Per Subject Per Studyid**
  - Leave study values as-is
  - TRTSDT/TRTEDT not useful for slotting
  - Subject basically treated as multiple unless summary programs “do the dancing” (not one proc-away!!)

- **Hybrid Approach**
  - One record per subject row
  - Individual study value row
  - Add a flag to identify the single record per subject row (UADSLFL)
Decision Based On?

• **Analysis needs**
  - Will the same subject’s data be summarized across studies?
  - How do I decide which AGE value to summarize?
  - How much subject data will be analyzed
    • Placebo to Active subjects
  - Will cross-study data be harmonized on an SDTM level?
  - Clear requirements and rules in the SAP are essential
ADSL Variable Values

- **Traceability considerations**
- **Change in place vs. creating new variables**
  - SDTM does not allow for “integrated” variables
  - Should ADaM?
Update or Create?

- **Update Existing Variables**
  - Standard is consistent
  - Original value and source gets lost
    - AGE recalculated based on a different anchor date

- **Create New Variables**
  - Original values maintained
  - New variables do not exist in the current standard
  - Programs have to be modified for new variables
Update or Create - AGE

- **Study 1** - based on randomization date
- **Study 2** - based on enrollment date
- **Study 3** - based on first dose date
- **ISS requires AGE based on first dose date**
  - Create “AGENEW” with consistent formula?
  - Create AGE with consistent formula?
    - Save original AGE as “AGEOLD” for traceability?
Treatment Variables

• **ISS – Low vs. High vs. Placebo**

• **Study 1 – High vs. Placebo**
  – Built with the end in mind: TRT01PN = 1 vs. 3
    • Breaks functionality of having TRTxxPN define column order
  – Study level: TRT01PN = 1 vs. 2
    • Requires recoding on an integrated level

• **Study 2 – High vs. Low vs. Placebo**
  – TRT01PN = 1 vs. 2 vs. 3
• Study 1: 30 mg vs. 60 mg :TRT01PN = 1 vs. 2
• Study 2: 20 mg Fed vs. 20 mg Fasted :TRT01PN = 1 vs. 2
• Study 3: 20 mg :TRT01PN = 1
• How does this get harmonized on an integrated level?
  - Create new treatment grouping variables?
  - Recreate TRTxxP based on integrated analysis needs?
• How do you handle subjects who get treated with multiple drugs if they are in the denominator for more than one column?
Grouping Variables

- Median value requires recalculation
- Active vs. placebo can have different meaning on an ISS level vs. study level
- Dose groupings differ for integration
Some More Recalculated Variables (example 1)

- **Study 1** - baseline is the average of Screening and Day 1 values
- **Study 2** - baseline is nominal Day 1 value
- **ISS** - needs harmonized definition - last non-missing value prior to dosing
Some More Recalculated Variables (example 2)

- **Study 1** - randomized trial - average daily dose based on titration period and then steady state dosing
- **Study 2** - open label study - average daily dose based on dosing log
- **ISS** - requires recalculation looking across both studies
Some Considerations

• Do we care what the study level value was in the ISS/ISE?
• Is there an expectation that study level analyses can be reproduced from the integrated datasets?
• If integrated ADaM is based on integrated SDTM do we still want study level ADaM values?
• What level of traceability do you want to provide?
Summary

- Decide where to start
- Gather all source data
- Decide on Structure (number of ADSL/obs)
- Update/Create New Variables
- Recalculate Necessary Variables
- Traceability
Questions?

• Nate Freimark - nate.freimark@TheoremClinical.com