Strategies for Preparing and Validating Submission Data

Sandeep Juneja, SAS Institute, Inc. / Frank Roediger, SAS Institute, Inc.

Everyone making clinical trial submissions wants the process to go as smoothly as possible, but there are some places where missteps can cause a subsequent delay in the process. Avoiding these missteps can greatly streamline the submission preparation process.

The CDISC SDTM Implementation Guides have hundreds of pages that provide exhaustive (and exhausting) detail about every imaginable domain and variable, but those Implementation Guides don't provide any information about how long any of the variables should be. Careful data design and some utility processes can take the guesswork out of declaring variable lengths.

V5 transport files are the approved mechanism for transmitting clinical trial data within a submission. Some of the archaic limitations of V5 transport files (for example, 8-character variable and format names) can now be overcome thanks to special-purpose macros that can be downloaded from SAS.

The define.xml contains information about what a reviewer can expect to find in a submission. Because discrepancies between the define.xml and the submission data sets can cause delays in the FDA review, it is very important to make sure that the define.xml truly reflects the submission data's metadata. The SAS Clinical Standards Toolkit (CST) provides a way to reconcile the define.xml with submission data metadata so that any discrepancies can be resolved before a submission is made.

Eight steps to manage the lengths of domain fields, generate .xpt transport files and define.xml, and verify that there are no metadata discrepancies between the .xpt and define.xml files.

1. **Individual studies:** domains are created with $200 character fields to avoid truncations
2. **Warehouse/Ocean:** domains still have $200 character fields (compression saves space)
3. **Mart/Pool:** domains' character fields are reduced to their values' maximum length
4. **Submission Preparation:** v5 .xpt files and define.xml
5. **Xpt/Xml Extracts:** metadata from .xpt files and define.xml
6. **Xpt/Xml Analysis:** compare metadata extracts
7. **Reconciliation Reports:** identify and address metadata discrepancies
8. **Submission-Ready:** v5 .xpt files and define.xml are ready to submit

References

- *Some Strategies for Validating Your Data before Submission*, Sandeep Juneja and Frank Roediger, SAS Institute, Inc., Cary, North Carolina. PharmaSUG 2013 Proceedings, May 12-15, 2013.
- http://support.sas.com/kb/46/944.html. Usage Note 46944: *New SAS transport format and tools available*
- http://support.sas.com/rnd/base/cdisc/cst/index.html - Clinical Standards Toolkit Information
- https://communities.sas.com/docs/DOC-1463 - Electronic copy of poster and supplemental handout

Contact Information

- SAS Drug Development Forum - https://communities.sas.com/community/support-communities/sas-drug-development
- Sandeep Juneja, Sandeep.Juneja@sas.com / Frank Roediger, Frank.Roediger@sas.com