Introduction

It is a constant struggle for sponsors to establish an integration strategy for standard (CDISC compliant) and non-standard data at SDTM and/or ADaM level, as no specific strategy or requirements exist within the industry and from regulatory agencies. The industry continues to face challenges and hurdles with posting, linking, traceability and compliance checks.

While there are a variety of scenarios that will arise during the preparation of the integrated databases in support of the ISS and ISE, this white paper is created by PhUSE SDTM and ADaM Implementation FAQ team to address frequently asked questions on guidance on how to prepare ISS/ISE, with the focus to be on two most common scenarios; note that any scenario used will require upfront discussions with the FDA and PMDA.

Scenario 1: ADaM-only Integration

ADaM-only integration strategy may be more suitable when:

- The majority or all of the IIS and ISE analyses and resulting tables, figures and listings (TFLs) can be generated using the ADaM datasets
- The ADaM datasets robust and complete enough to be used to conduct all the analyses required by the IIS and ISE
- A very small number of the studies selected to be included in the integration analysis are not CDISC complaint with SDTM standards
- There are little to no new derivation needed in the IIS and ISE analyses
- There is minimal dictionary or controlled terminology aversion/required
- There was sufficient pre-planning for integration included in the ADaM datasets

Not all data that is required for the IIS and ISE analyses is included in the study ADaM datasets and it is necessary to obtain it from the study SDTM datasets

Study SDTM and/or Legacy Data \(\rightarrow\) Integrated ADaM

Scenario 2: SDTM + ADaM Integration

Sponsor may choose to integrate both the SDTM and the ADaM databases when:

- The majority of the studies to be included in the integration have CDISC-compliant SDTM and ADaM databases in place
- There is a large mixture of studies with standard and non-standard study data to be included in the integrated analyses
- There are too many different standards versions , controlled terminology versions, or dictionary versions that need to be harmonized
- The number of IIS and ISE TFLs that can use SDTM as the data source is significant
- Not all the data that is required for the IIS and ISE analyses is included in the study SDTM datasets

Integration needs extend beyond IIS and ISE analyses and are required for other tasks

Study SDTM \(\rightarrow\) Integrated SDTM \(\rightarrow\) Integrated ADaM

Adherence to Regulatory Agency Guidance

Up-versioning of Controlled Terminology (CT)

- The SDTM CT dictionary to specifically discuss the use of standardized and custom controlled terminology, and may be different for the new SDTM datasets.
- Note: up-versioning of standardized CT is not mandatory for the within individual SDTM or ADaM dataset.
- Recommended to harmonize the CT (most current version available at the time) during integration activities.
- Refer to PhUSE white paper on CT versions for different scenarios.

Harmonization of Coding Data

- FDA and PMDA recommend to utilize MeDRDA to code adverse events (AE) and WHODrug to code adverse events (AE) and WHODrug.
- AE and CMs should be coded to a single version of the respective dictionaries in the IIS and the most recent versions should be used at the time that the studies are pooled.
- When harmonizing coding dictionaries like MeDRDA and WHODrug, the IIS has recommended to provide a table that lists all events whose preferred term or hierarchy has changed when the data was converted from one dictionary version to another.

Define.xml and DRG Generation

- The define.xml and DRG needs to be created for the new integrated datasets.
- For scenario 2, provide these documents for both databases if both are included in the submission.
- Harmonizing of coding dictionaries provide a table of converted data from one dictionary version to another.

Traceability and Compliance Considerations

- Scenario 1 may require the generation of intermediary datasets for dictionary, CT, or standards version harmonization.
- These intermediary datasets will not be submitted to FDA or PMDA
- Scenario 2 provides a clear, traceable path from study data to integrated ADaM datasets which eliminates the need to generate intermediary datasets.
- Each step of the process (study level data, integrated SDTM data, integrated ADaM data) is submitted to FDA and PMDA.
- Variables - e.g. xxSEQ, ASEQ, SRCH, SRCDOM, ASEQ - that are generally used in study level SDTM and ADaM datasets to depict traceability may not be applicable in integrated datasets.
- Scenario 1: xxSEQ can be used to show traceability to study level data.
- Scenario 2: xxSEQ will need to be regenerated in the integrated SDTM datasets and will not provide traceability to study level data.
- Recommended to run P21 checks, even though FDA and PMDA validation rules are written for study SDTM data and study ADaM data.
- Scenario 2 may generate more compliance errors/warnings than Scenario 1, one set for integrated SDTM data and another set for integrated ADaM data.

Conclusion

The whitepaper is a work in progress and is aimed to be published in Q3 2019. We are always looking for more members! Contact the author to join us.

References


Acknowledgments

We extend our thanks to all the members of the PhUSE SDTM and ADaM Implementation FAQ Project Team.