Paper SI08

Case Study: Integrated SDTM to Integrated ADaM to TLFs

Kishore Pothuri, Vita Data Sciences, Waltham, MA, USA
Kesava Vajjala, Vita Data Sciences, Waltham, MA, USA

ABSTRACT

The Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE) are essential components of a successful submission. It can be very challenging and tedious to pool data for analysis from multiple studies without losing traceability, since similar information could have been mapped differently in different studies through diverse systems by various vendors, thereby putting compliance and harmonization at risk. What complicates the task further is that the number of studies involved can range anywhere from 2 to 50, or even more. There are multiple approaches to pooling and integrating study datasets for ISS and ISE analyses, namely ADaM Integration using study-level SDTM domains, ADaM Integration using Pooled SDTM datasets, ADaM Integration using study-level ADaM datasets. This paper mainly focuses on ADaM Integration using Pooled SDTM datasets approach and discusses in detail about the challenges faced, do's and don'ts, lessons learnt and the preemptive measures which can be taken to avoid any hurdles to efficiently program and standardize the data when pooling SDTM datasets.

INTRODUCTION

Definition:

Integrated summary of safety and integrated summary of efficacy are not just summaries, as the name might suggest, but rather documents comprised of integrated analyses of the safety and effectiveness of a study drug. In other words, the results from all clinical trials performed on the study drug are pooled together and analyzed as a whole, producing combined statistical results [1].

ISS and ISE: Needed for a regulatory submission?

ISS and ISE are vital aspects of New Drug Applications (NDAs), uniquely required by FDA (USA) regulation [2]. These integrated analyses are not strictly required for NDA submissions to the MHLW (Japan) and the EMA (EU), however MHLW and EMA submissions must follow the common technical document (CTD) format, which contains sections in line with ISS and ISE [3]. As of January 1st, 2018, the CTD format is also mandatory for certain regulatory submissions to Health Canada including NDAs, referred to as a New Drug Submission (NDS) by Health Canada [4]. Both ISS/ISE and the sections of the CTD have essentially the same requirements, however the ISS and ISE are more rigidly formatted and require a great amount of detail of the analysis and supporting information. These integrated analyses are not required for biologic license applications (BLAs), however it is strongly recommended that they are included.

Whether for regulatory submission or not, integrated summaries are useful for:

- Discovering rare and/or unexpected trends in patient subgroups
- Improving the precision of results with a larger population size by integrating studies
- Comparing variation in study results to assess the risk and benefits
- Reaching a strong and defendable statistical conclusion

Data Pooling and Integration Approaches:

As most of us are aware, there isn’t an established process within the clinical industry, or any proven strategy defined by regulatory agencies for pooled data analysis which involve multiple studies. Hence, it is a continuous challenge towards which a lot of effort is being put in, to meet this need by many sponsors and vendor companies. Some of the popular approaches for data pooling and integrating study datasets for ISS and ISE analyses includes:

ADaM Integration using study-level Tabulation datasets:

The figure 1 below shows the elements involved in the approach using Tabulation Datasets (whether they are SDTM or Raw datasets) to create integrated ADaMs.

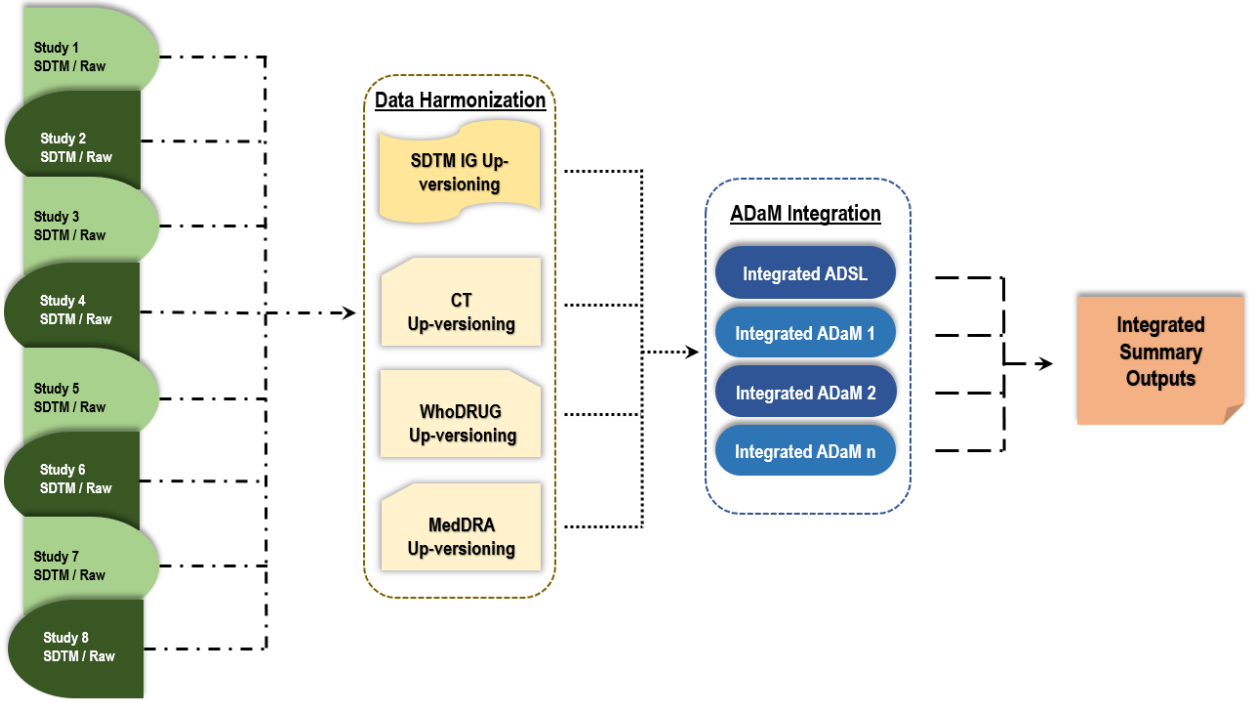

Figure 1: ADaM Integration using study-level Tabulation datasets

This strategy may be applicable when any of the below scenarios are met [5]:

- If any or some of the studies that are part of integration are not CDISC compliant with respect to SDTM mapping
- If there a blend of studies with standard and non-standard study data included in the integrated analyses

ADaM Integration using study-level Analysis datasets:

The figure 2 below represents the approach using study level ADaM datasets for ADaM Integration.

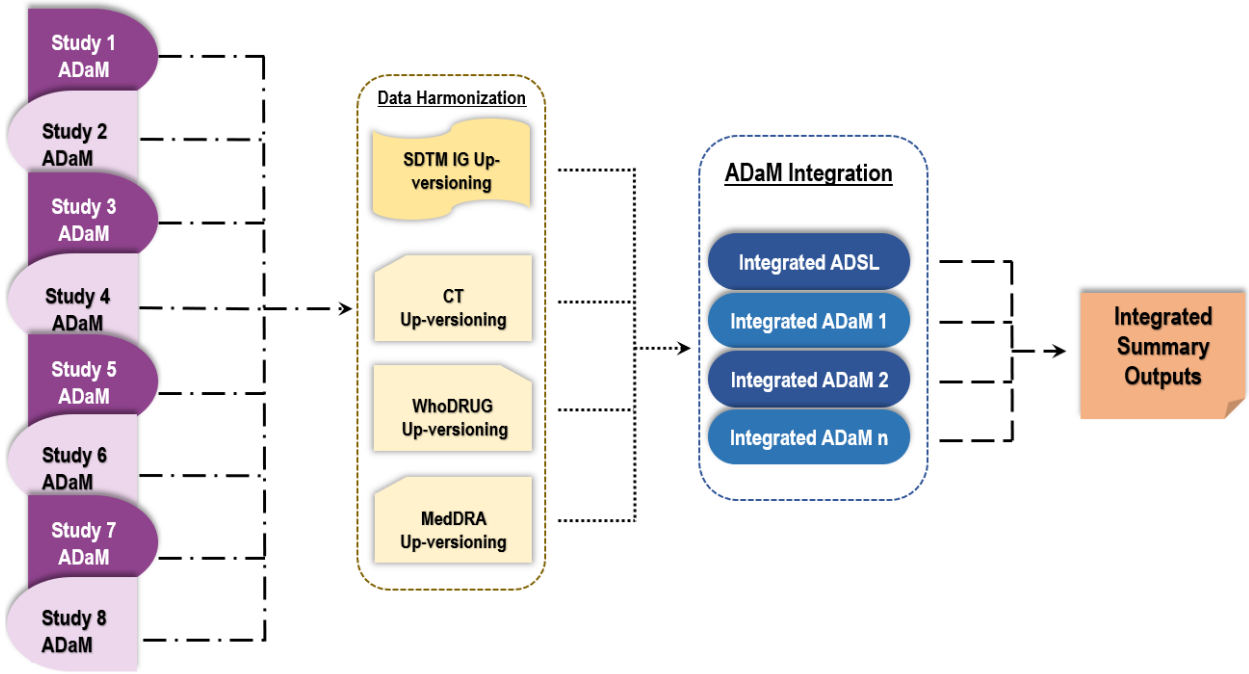

Figure 2: ADaM Integration using study-level Analysis datasets

This strategy may be applicable when any of the below scenarios are met [5]:

- If any or some of the studies that are part of integration are not CDISC compliant with respect to SDTM mapping
- If the outputs (Tables, Listings & Figures) for ISS and ISE analyses can be developed solely from the ADaM datasets
- If the ADaM datasets are robust and complete enough to conduct all analyses required by ISS and ISE

ADaM Integration using pooled SDTM datasets:

In this paper, we will discuss in detail about a case study where the 'ADaM Integration using Pooled SDTM datasets approach' was followed and highlights some of the challenges faced, pros & cons and how this approach helped in the ease of ADaM datasets programming and in generation of Tables, Listings and Figures.

CASE STUDY: INTEGRATED SUMMARY OF SAFETY ANALYSIS USING POOLED SDTM DATASETS

The figure 3 below depicts the approach adopted in the case study to use study level SDTM datasets to create pooled SDTMs which were then used as a source to create Integrated ADaMs.

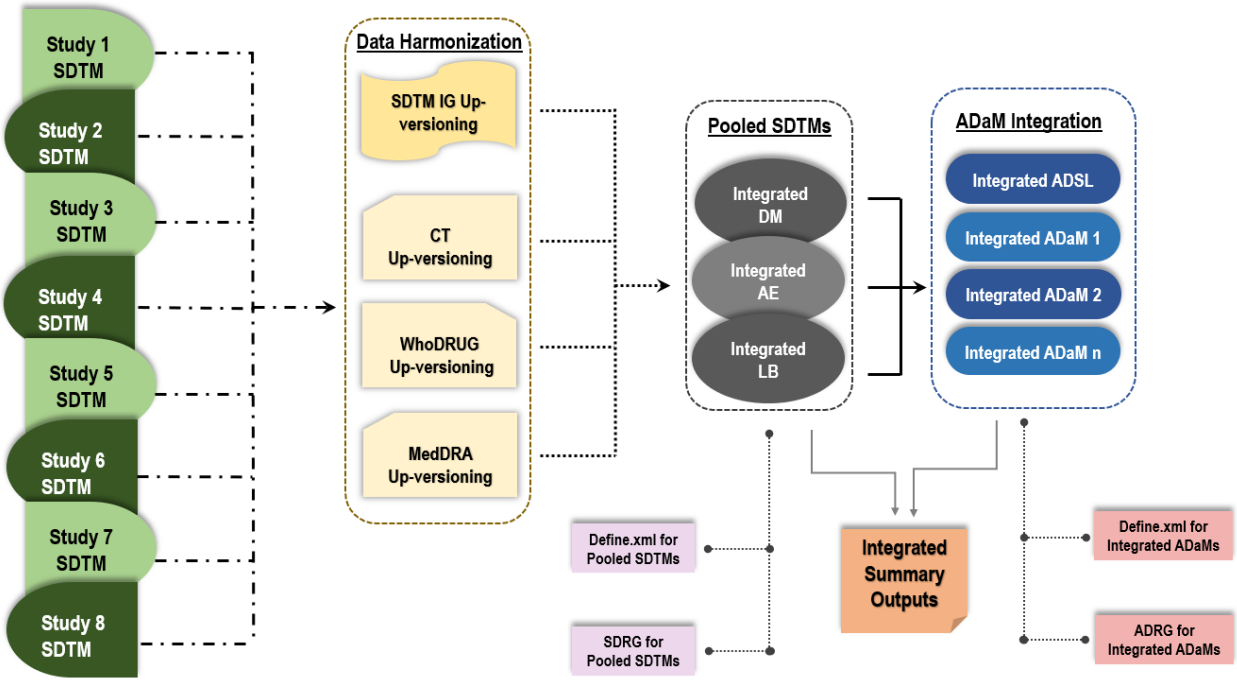

Figure 3: ADaM Integration using pooled SDTM datasets

RATIONALE:

The task was to pool data from 8 studies for integrated summary of safety (ISS) analysis for a regulatory submission and for this purpose, the study-level SDTM datasets were used to create pooled SDTM datasets to generate the integrated ADaM datasets, which were then used as a source to create the Tables, Listings and Figures. Below are the main driving factors in the decision to choose this approach over the others:

- Majority of the studies which were part of data pooling were done by the same vendor (which is VDS in our case) and were already CDISC compliant, ranging from the year 2014 through 2019.

- The therapeutic area was 'Rare Diseases' and hence the number of studies involved were less and facilitated the stacking the SDTM datasets from multiple studies. The table below provides details on each of the studies.

Study	SDTM IG Version used	SDTM CT Version used	MedDRA Version used	WHO Drug Version used	Vendor
STUDY 1	3.1.3	2014-03-28	17	WHO (World Health Organization) Drug Version September 2014 Enhanced Dictionary Version B2	VDS*
STUDY 2	3.2	2017-12-22	18.1	WHO Drug Dictionary Q2 2015	VDS*
STUDY 3	3.2	2017-12-22	18.1	WHO Drug Dictionary Q2 2015	VDS*
STUDY 4	3.1.3	2014-10-06	17	WHO (World Health Organization) Drug Version September 2014 Enhanced Dictionary Version B2	VDS*
STUDY 5	3.1.3	2014-06-27	17	WHO (World Health Organization) Drug Version September 2014 Enhanced Dictionary Version B2	VENDOR 1
STUDY 6	3.2	2017-12-22	20.1	WHODrug B3 Sep 2017	VDS*
STUDY 7	3.2	2017-12-22	21	WHODRUG B3 GLOBAL (MARCH 2018)	VDS*
STUDY 8	3.2	2018-12-21	21	WHO Drug Mar 2018G B3	VENDOR 2

***VDS= Vita Data Sciences**

- The approach provided ease of ADaM datasets development from a single source of pooled SDTM datasets, rather than creating ADaM datasets from individual SDTM datasets from multiple studies.
- Initially, the efforts with this approach were estimated to be about 10% more than the efforts needed for developing integrated ADaM datasets using study level ADaM datasets from multiple studies. However, at the end we were able to do it much quicker than we anticipated. Moreover, this was still the chosen method because 3 out of 8 studies were still ongoing and at the long run whenever a new data cut is taken for the ongoing studies, only the pooled SDTM datasets will need an update and there is absolutely no necessity to update ADaM dataset programming.
- Maintenance of ADaM programming and datasets is more tedious than SDTM, which can be avoided with this approach.

The idea was to harmonize the mapping at pooled SDTM level rather than at integrated ADaM level. The figure below explains the process of Data Harmonization at the pooled SDTM level.

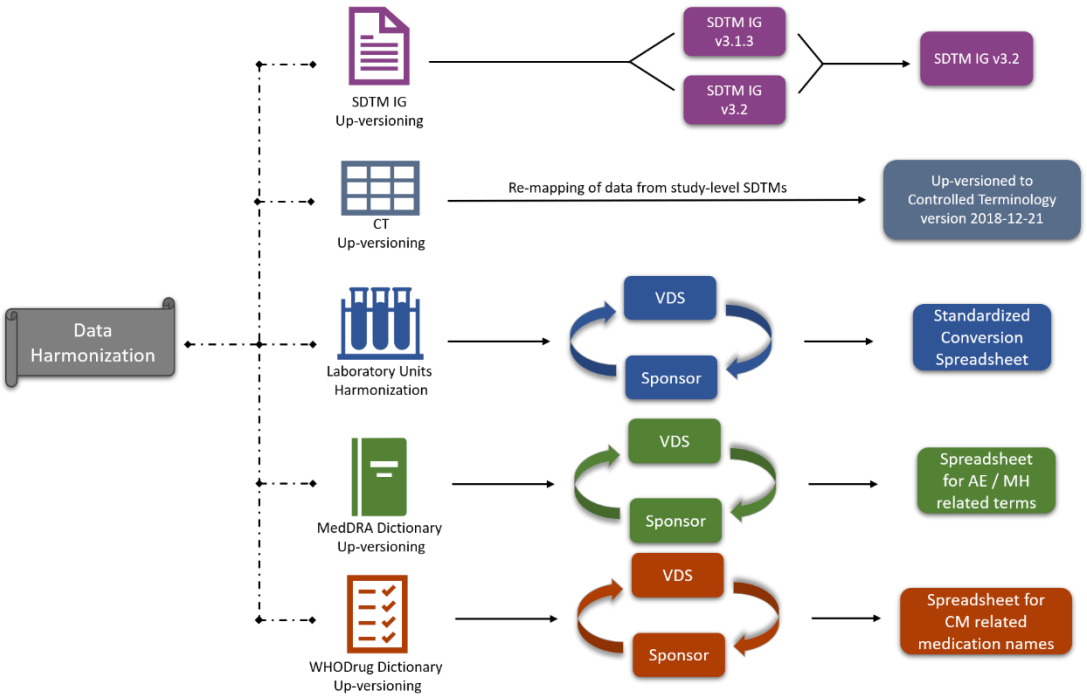

Figure 4: Data harmonization at the pooled SDTM level

All the studies were up versioned to use SDTM IG v3.2. This required minor updates as most of the studies were CDISC compliant. To harmonize the Controlled Terminology, some information was re-mapped into the appropriate domains to facilitate the stacking of SDTM datasets. The tricky part, however, was to deal with the up versioning of the MedDRA and WHODrug dictionaries as well as harmonizing the Laboratory test units. The fact that few of the

studies were still ongoing meant the addition new unit values each time the updated eDC was made available. To tackle this, spreadsheets were created which comprised of all the unique Adverse Event/Medical history terms, Concomitant Medication Treatment names and the units pertaining to each laboratory test. These spreadsheets were then sent over to the sponsor, who in turn provided the appropriate values to which the term or unit needed to be up versioned to. The spreadsheets were then incorporated into the respective domains at the pooled SDTM level to populate the values as needed.

CHALLENGES:

As with any novel or innovative approach, there are corresponding challenges associated. Below are some of the challenges we faced:

- During this project, there were no reference studies or similar processes followed.
- Handling roll over subjects for some of the studies which were part of this integration.
- Some studies were legacy and had slightly different SDTM mapping approaches although compliance was not a hurdle.

PROS AND CONS OF THE ADAM INTEGRATION USING POOLED SDTM APPROACH:

The following sections discuss some of the pros and cons of implementing this approach over the others and highlights the lessons learnt from the case study mentioned above.

PROS	CONS
• The need for study-level ADaM datasets was eliminated, as all the information is available at the pooled SDTM datasets. This avoided much of the data transformation (such as merging data from multiple studies for instance) which is very typical when generating ADaM datasets. • Pooling the SDTM datasets from the studies proved to be an easy task as majority of the variables were standardized and CDISC compliant at the study level. • The conversion and/or harmonization of the Lab units was done at the pooled SDTM level rather than the integrated ADaM level. This reduced the time and effort that went into generating and validating the analysis Lab dataset, ADLB, which is a complex dataset on its own, let alone an integrated ADLB dataset. • Up-versioning of the MedDRA and WHODrug dictionaries was much easier at the pooled SDTM level rather than the individual study level SDTMs. Similarly, harmonizing the Controlled terminologies (CT) at the pooled level was much more efficient rather than doing the same at the individual study level.	• Deriving new variables could not be done at the pooled SDTM level; it was more feasible to do the same at the integrated ADaM level. This complicated the derivation slightly as we had study-specific information required from multiple pooled SDTMs. • Standardizing or harmonizing the Visit information – VISIT, VISITNUM and other variables such as --TPT (Planned time point name), --TPTNUM (Planned time point number) values, etc. at the pooled SDTM level proved to be complicated due to the varying conventions from the individual studies. • In some cases, the traceability factor was impacted due to mapping decisions taken at an individual study level. This issue was further compounded as couple of the studies were done by an external vendor; this resulted in non-uniform mapping of the information across some of the SDTMs. For instance, information related to dose interruption, etc. was mostly collected in the EX, SUPPEX, EC and SUPPEC domains, but one study collected this information in the FA domain. • Validation of the pooled SDTM and integrated ADaM datasets proved to be slightly time consuming as the pinnacle 21 validator tool was not an ideal platform for such a task. The issue report generated was lengthy due to the stacked nature of the datasets and thus was time taking when the said issues required an explanation or justification.

LESSONS LEARNED:

The case study overall was a success. However, it came with a few lessons learnt. If the scope is having an ISS/ISE analysis as the end goal for a sponsor, it is in the best interest that the decisions related to mapping and standards are made in a uniform manner upfront when mapping data to the study level SDTMs. This translates into saving time and being highly efficient when creating pooled SDTM datasets. It is highly beneficial to have the same team working throughout the integration process to avoid any misinterpretation of data, and to facilitate consistent decision making. This also ensures steady functioning from an operational standpoint. This approach can be further streamlined if all the studies involved in the integration are handled by a single vendor.

CONCLUSION

The ADaM Integration using Pooled SDTM datasets approach had helped us in streamlining the ISS process. Additionally, it significantly reduced the amount of time that is typically needed to program ADaM datasets using individual SDTM domains when multiple studies are involved. This white paper is solely written based on the scenarios that we had encountered and doesn't necessarily impose any individual or organization to follow the approach mentioned in this paper.

REFERENCES

[1] https://www.quantics.co.uk/blog/an-introduction-to-integrated-summary-of-safety-and-integrated-summary-of-effectiveness-iss-and-ise/

[2] FDA Center for Drug and Biologics Evaluation and Research (CDER and CBER). (April 2009). *Guidance for Industry: Integrated Summaries of Effectiveness and Safety: Location Within the Common Technical Document*. Retrieved from https://www.fda.gov/downloads/drugs/guidances/ucm136174.pdf.

[3] (2018). *CTD: M4: The Common Technical Document*. [online] Retrieved from http://www.ich.org/products/ctd.html.

[4] ca. (2018). *Updated Notice – Mandatory use of the Electronic Common Technical Document (eCTD) format – Canada.ca.* [online] Retrieved from https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/activities/announcements/notice-mandatory-use-electronic-common-technical-document-ectd-format.html.

[5] ISS and ISE Dataset Preparation Best Practices: A PhUSE Whitepaper
Aatiya Zaidi, Kapila Patel, Bhavin Busa, Veena Nataraj, Kiran Kumar Kundarapu, Joanne Zhou, Soumya Rajesh
https://www.lexjansen.com/phuse-us/2019/pp/PP14_ppt.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Author Name: Kishore Pothuri
Company: Vita Data Sciences
Address: 281 Winter St, suite# 100
City / Postcode: Waltham, MA, 02451
Work Phone: 781-833-0259
Email: kpothuri@vitadatasciences.com
Web: www.vitadatasciences.com

Author Name: Kesava Vajjala
Company: Vita Data Sciences
Address: 281 Winter St, suite# 100
City / Postcode: Waltham, MA, 02451
Work Phone: 781-833-0259
Email: kvajjala@vitadatasciences.com
Web: www.vitadatasciences.com