Paper SA07

Navigating FDA requirement: ISS/ISE build strategies

Priyank H Patel, Accenture, Berwyn, USA

ABSTRACT

An Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE) are critical components of successful new drug applications (NDAs) and new biologics license applications (BLAs). These are not only summaries but enablers to marketing and licensing applications that provide detailed views into integrated safety and efficacy profiles across all relevant studies within an application. In order to gain a more complete analytical picture, we frequently see legacy, or non-CDISC studies as part of these integrated summaries. These non-standardized data can make the overall study-level integration activities tedious, costly and time-critical. A key submission rejection criterion from regulatory authorities such as Food and Drug Administration (FDA) is the failure to define the traceability of the data from its collection through to clinical study reporting. In this presentation we will explore multiple options within an ISS/ISE strategy to handle legacy study data, overall integration planning and assurance of traceability as to mitigate lengthy review cycles or refuse to file outcomes.

INTRODUCTION

An Integration Analysis is the standard method to collectively describe the Effectiveness and Safety of a study drug in development. The summaries are intended to provide a more transparent understanding of responses across different populations (demographics, disease-related, etc.), and dosing regimens. The particulars of the integration analysis data summaries are highlighted in the clinical study report (CSR) to illustrate a better perspective on drug performance across multiple trials. It specifically requires a thorough planning if the new drug and biologics applications include studies that are in the legacy format aka non-CDISC standard format. FDA expects to see all the clinical trials data adhering to CDISC format in the submission application. The non-CDISC data format may bring several unknown challenges, mainly tracing all key data information needed for the analysis. Also, it may delay the overall assessment process - especially, if the sponsor did not implement uniform data collection standards during the study build set-up. It becomes labor-intensive and cost-intensive. The effort increases based on the size, and the number of legacy studies that are planned to be included in the submission. There are several ways of planning an integration analysis, but the key is attentive planning from inception. The determining factor is if the data flow can demonstrate traceability from the Case Report Form (CRF) to the CSR organically, and shouldn't be forced.

LEGACY STUDY DATA HANDLING PRIOR TO INTEGRATION

The first thing to recognize is that FDA does not support study data in a non-CDISC format as mentioned in the Study Data TCG of the section 8.3.2 [(1)]. It is expected that the sponsor should convert data from the legacy format to CDISC format by mapping every data element originally collected to a respective data element of the CDISC standard. During this transformation, there may be cases of simple data migration from a legacy to a standardized format. Also, there may be cases of where the data transformation can be complex, and the legacy data cannot be fully standardized. In these cases, the expectation is that the rationales of these conversions should be provided in the Study Data Reviewer's Guide (SDRG).

Primarily, if the start date of the legacy study is prior to December 17, 2016, FDA may allow a sponsor to submit the data in the original legacy format (i.e. legacy aCRF, legacy tabulation data, legacy analysis data, and legacy analysis programs) in conjunction with the CDISC SDTM tabulation data package at the minimum to support FDA's review process. Again, FDA does not recommend a particular approach to legacy clinical study data conversion as mentioned in the section 8.3.2.1 of the Study Data TCG, but expects that the sponsor should have a plan in place to address them to ensure the converted data are

traceable and adequate to support the review process [1]. The data standard catalog provides the guidelines to select the standard that FDA accepts as shown in Table 1 below [2].

FDA Data Standards Catalog v6.1 (09-09-2019) - Supported and Required Standards

For full description of column headings, see Instr.& Column Descriptions tab

Use	Data Exchange Standard	Exchange Format	Standards Development Organization (SD	Supported Version	Supported Implementation Guide Versio	FDA Center(s)	Date Support Begins (MM/DD/YYYY)	Date Support Ends (MM/DD/YYYY)	Date Requirement Begins (MM/DD/YYY	Date Requirement Ends	Statutory, Regulatory, or Guidance Authority	Information Sources
Clinical study datasets	Study Data Tabulation Model (SDTM)	XPT	Clinical Data Interchange Standards	1.1	3.1.1	CDER, CBER	Ongoing	01/28/2015			Standardized Study Data	CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	3.1.2	CDER, CBER	10/30/2009	03/15/2019 [1] 03/15/2020 [2]	12/17/2016 [1] 12/17/2017 [2]	03/15/2019 [1] 03/15/2020 [2]	Standardized Study Data	CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.2	Version 3.1.2 Amendment 1	CDER, CBER	08/07/2013	03/15/2019 [1] 03/15/2020 [2]	12/17/2016 [1] 12/17/2017 [2]	03/15/2019 [1] 03/15/2020 [2]	Standardized Study Data	CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.3	3.1.3	CDER, CBER	12/01/2012		12/17/2016 [1] 12/17/2017 [2]		Standardized Study Data	CDISC.org - SDTM
Clinical study datasets	SDTM	XPT	CDISC	1.4	3.2	CDER, CBER	08/17/2015		03/15/2018 [1] 03/15/2019 [2]		Standardized Study Data	CDISC.org - SDTM
Clinical study datasets	Analysis Data Model (ADaM)	XPT	CDISC	2.1	1.0	CDER, CBER	Ongoing	03/15/2019 [1] 03/15/2020 [2]	12/17/2016 [1] 12/17/2017 [2]	03/15/2019 [1] 03/15/2020 [2]	Standardized Study Data	CDISC.org - ADaM
Clinical study datasets	ADaM	XPT	CDISC	2.1	1.1	CDER, CBER	03/15/2018		03/15/2019 [1] 03/15/2020 [2]		Standardized Study Data	CDISC.org - ADaM

Table 1 – FDA Data Standards Catalog v6.1 [2]

FDA VIEWPOINT:

It is important to understand FDA's expectations and their review process. Figure 1 shows a high-level overview of FDA review process. The set-up shown in the process helps the reviewer perform a transparent, consistent, and efficient scientific review by performing automated extraction, transformation, loading, management, and integration of data to facilitate regulatory review [3]. FDA does not accept any clinical trial data in non-CDISC format unless a waiver has been granted.

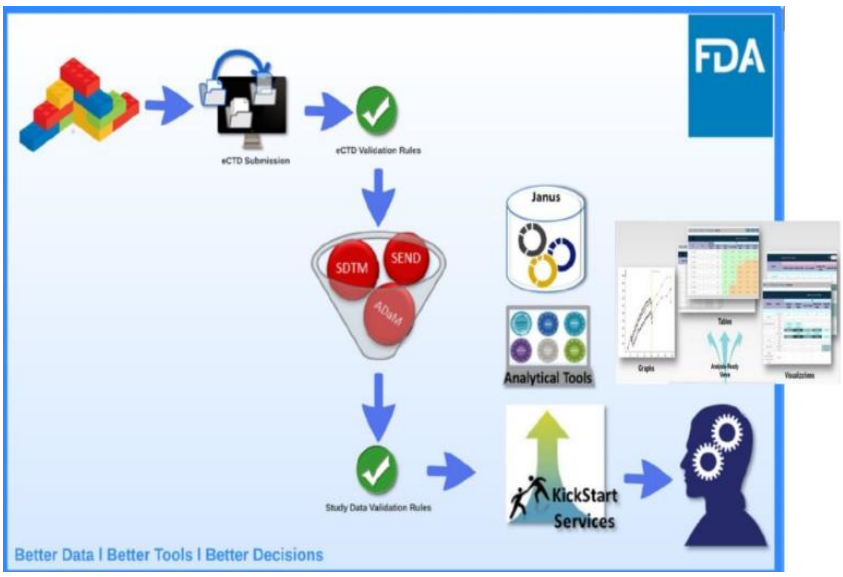


Figure 1 – FDA Review Approach [3]

FDA expects the sponsor must shadow the guidelines provided by CBER and CDER for respective applications. This includes following the data standards which allow the FDA to modernize and streamline the overall review process. The data standards facilitates to exchange of clinical research data between different computer systems and provide a general structure for organizing study data, including templates for datasets, standard names for variables, and standard ways of doing calculations with common variables [4].

The Study Data TCG is a key to understand FDA's current thinking (Figure 2) on the regulatory submission approach in electronic format. If one must use an alternate approach, one should consult FDA staff responsible for implementation of this guidance. As it stated "This Study Data Technical Conformance Guide (Guide) provides specifications, recommendations, and general considerations on how to submit standardized study data using FDA-supported data standards located in the FDA Data Standards Catalog (Catalog). The Guide supplements the guidance for industry Providing Regulatory Submissions in Electronic Format — Standardized Study Data (eStudy Data). The eStudy Data guidance implements 745A(a) of the Food, Drug, & Cosmetic (FD&C) Act with respect to standardized study data contained in certain investigational new drug applications (INDs), new drug applications (NDAs); abbreviated new drug applications (ANDAs); and certain biologics license applications (BLAs) that are submitted to the Center for Drug Evaluation and Research (CDER) or the Center for Biologics Evaluation and Research (CBER)." [1].

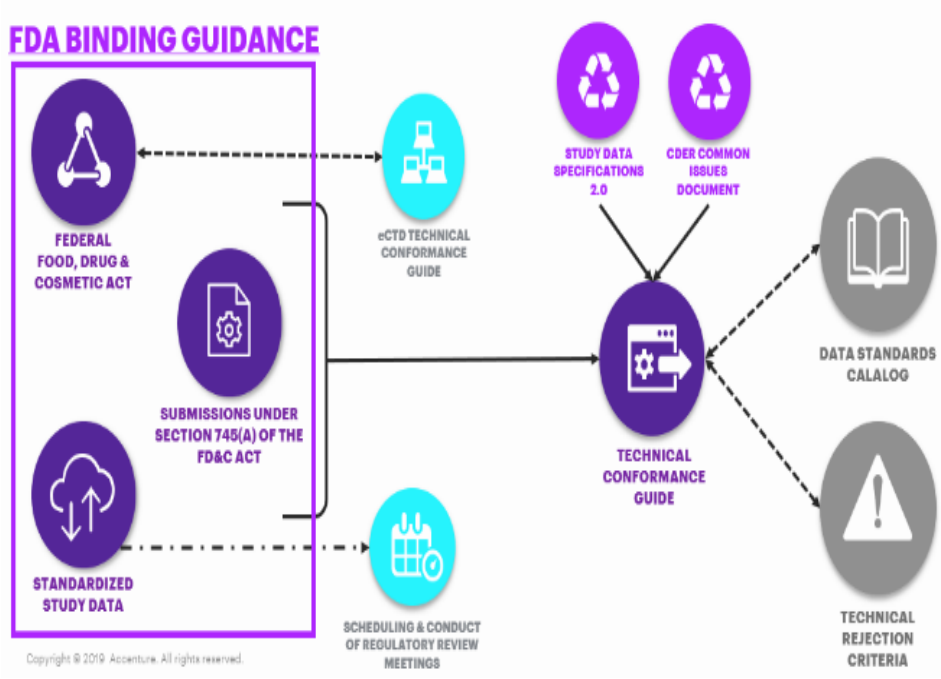


Figure 2 – Resources Navigating FDA Requirement

INTEGRATED ANALYSIS PLANNING:

FDA expects the sponsor to submit trial data using CDISC standard format for any application that involves an ISS, and ISE. As shown in Figure 3, there are various possible options to strategize an integration analysis approach. There are pros and cons to each option. Option 1 is based on performing an integration analysis data pool using ADaM data of individual studies. Option 2 is based on developing an integration analysis data pool directly from SDTM data of individual studies. In option 1 and 2, the standardization effort can become extensive and get muddled due to the involvement of substantial programming effort. It can require extensive documentation to explain the data transformation and standardization to prevent any confusion on the reviewer's end. Option 3 is based on creating an integrated SDTM pool from individual studies' SDTM data package. This option provides the sponsor higher flexibility to attain standardized and harmonized data. It also helps in defining complete traceability to integrated ADaM datasets and TLFs as the integrated SDTM pool data can be a single source of input downstream.

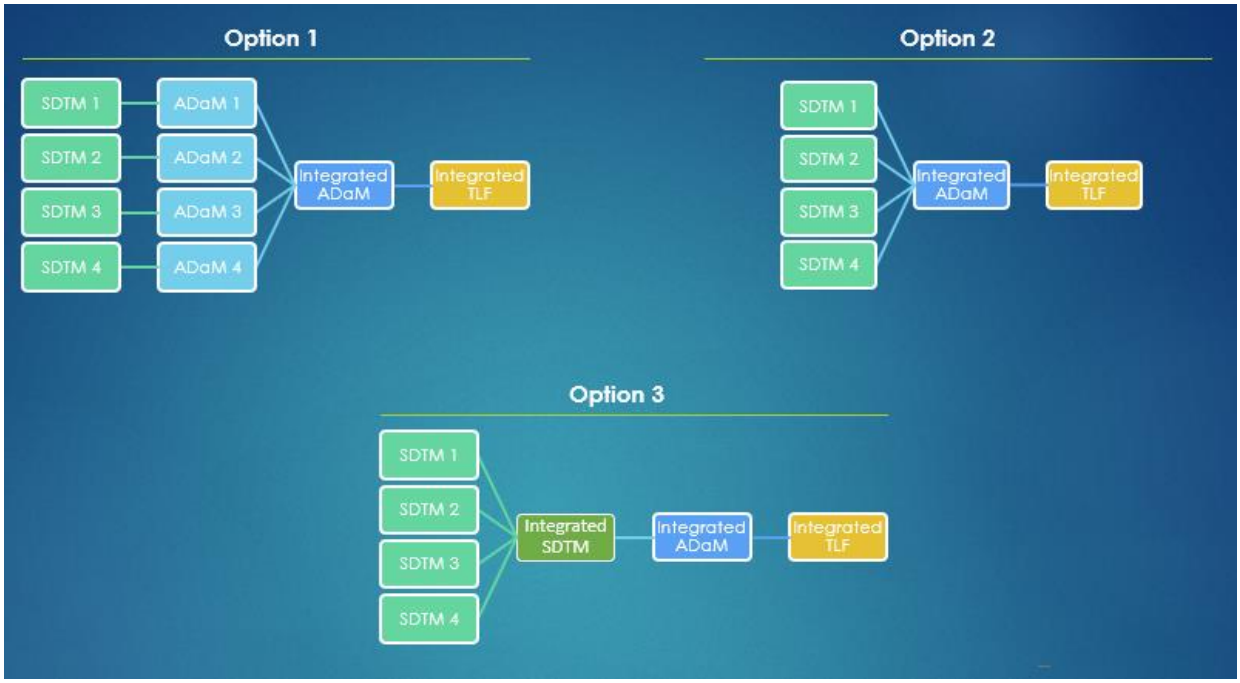


Figure 3 – Integration Analysis Possibilities

Options 3 allows the following but, not limited to:

1. Seamless SDTM presentation
2. USUBJID harmonization
3. Drug dictionary and Controlled Terminology standardization
4. Key analysis parameters standardization
5. Dataset splitting consistency
6. Opportunity to rationalize the data flow and related concerns into the publishing documents proactively

Seamless SDTM presentation:

The statistical analysis of integrated data is a complex procedure. It becomes a challenge if the studies are developed using different study data standards. Option 3 allows stabilizing the data transformation to one acceptable data standard accepted by FDA using a fluid and traceable approach as shown below in Figure 4.

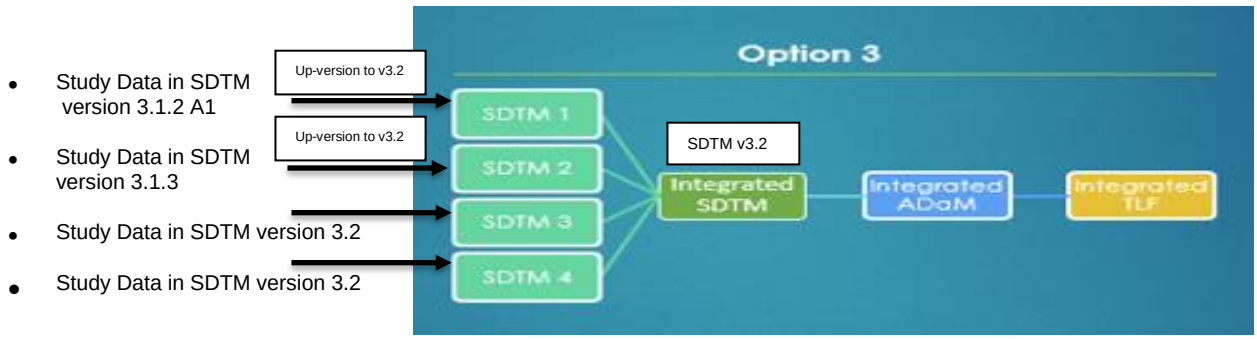


Figure 4 – Seamless SDTM development approach

USUBJID harmonization:

There may be cases where the same subject may have participated in more than one trial due to the nature of the study types or designs.

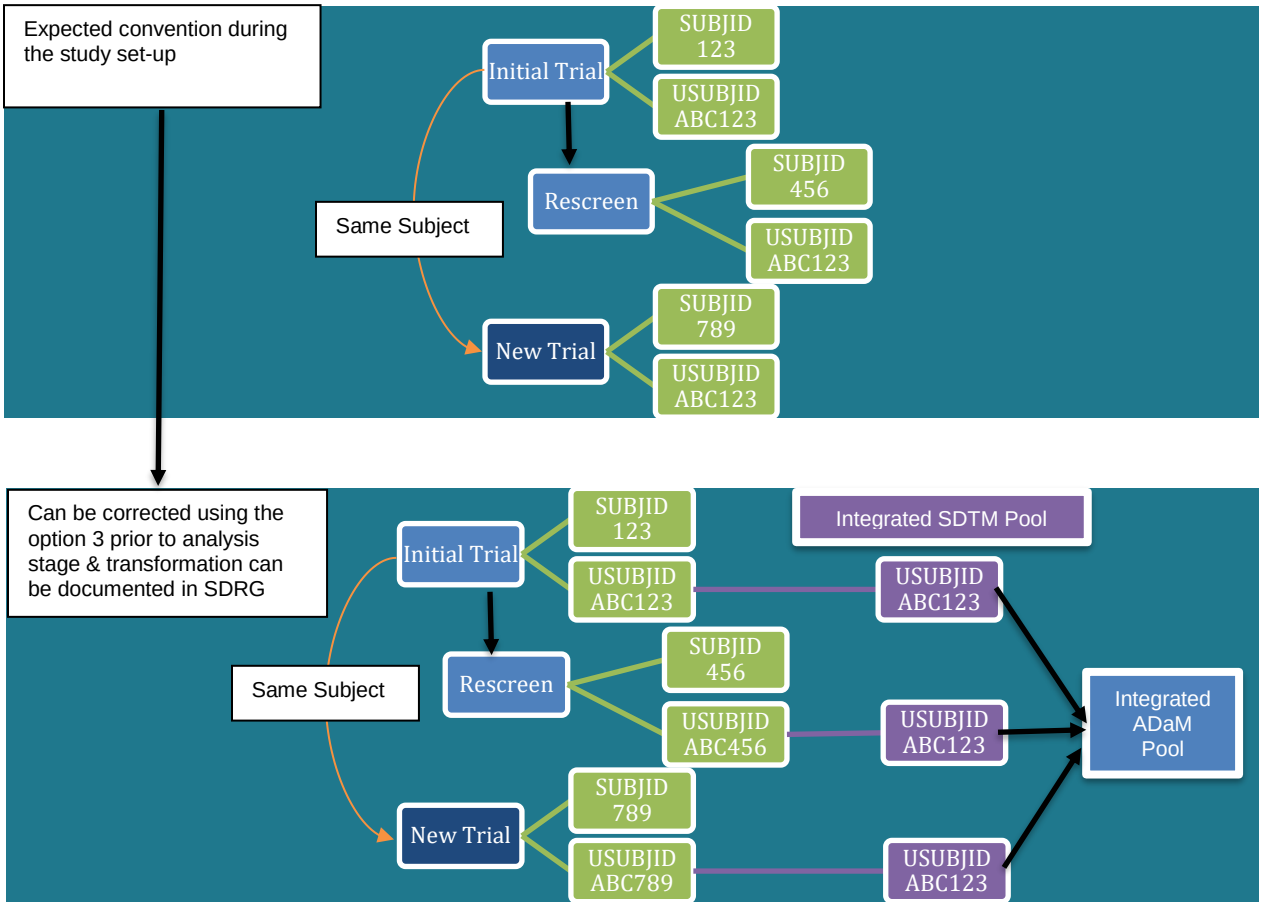


Figure 5 – USUBJID Harmonization Effort

The subject must have a unique identification for each trial. In CDISC SDTM world, the SUBJID is the individualized ID of the subject participating in a study, and the USUBJID is used to uniquely identify a subject across the submission portfolio. The sponsor should control the list of the unique trial subjects across portfolio which can support harmonizing and verify the USUBJID data information at the integration level as exemplified in Figure 5.

In some cases where studies were conducted in multiple capture systems, technologies, or there is some other novel difference in approach, it may be necessary to identify potential individuals who participated in multiple trials where it's not explicitly documented in the data. In this case, often, the demographic characteristics such as sex, age, birth date or birth year, race, ethnicity, along with vital characteristics such as height with other unique subject characteristics get analyzed to construct the decision-making process of identifying the USUBJID in absence of a valid evidence from the sites. It is recommended to discuss such findings with FDA as early as possible to get their feedback promptly.

Drug dictionary and Controlled Terminology standardization:

The Study Data TCG section 6.1 recommends that the common drug dictionaries such as MedDRA and WHODrug should be implemented across all clinical studies for each of the following: adverse events, medical history, concomitant medications, procedures. It also suggests using CDISC Controlled Terminology values for the standardization effort. The use of sponsor defined terminology or extension of a standard dictionary should be avoided. But, if necessary, the specifics should be documented in the define.xml, and the SDRG[1]. Different trials may have implemented different versions of Controlled Terminologies or drug dictionaries at a given point per the study requirements. All these inconsistencies should be fixed before the data gets considered final for the integration analysis use. Option 3 allows easier passage to implement the same standards across studies at once without changing the individual study specific requirements.

Key analysis parameters standardization:

If the trials' data conversion services are run through different vendors by a sponsor, chances of inconsistent display of key safety and efficacy information in tabulation format are higher. It can bring unknown challenges during the integration analysis effort. Option 3 allows unifying this key information at the integrated SDTM pool level by maintaining data traceability with proper documentation. The best way to conduct this effort is by using the look-up table approach. It allows a user to demonstrate a complete transformation of how data appears at the study level vs. integrated level.

The most common example of safety data is laboratory data. Laboratory data can originate from different local laboratories and central laboratories at a time for different trials, and it can bring significant challenges standardizing the lab parameters for the integration analysis purpose. The recommended approach is that the same lab parameter should have all results presented in one standardized unit per lab category shown in Table 2. This standardization effort is SAP driven, which gives an overview of the effort required to standardize the lab parameters across the trials at the integration level.

Source Information						SDTM Conversion (Study Level)									SDTM_Conversion (Pool Level)	
SRC_LBSPEC	SRC_LBCAT	SRC_LBTESTCD	SRC_LBTEST	SRC_ORRESU	SRC_STRESU	SDTM_LBTESTCD	SDTM_LBTEST	SDTM_LBCAT	SDTM_LBSPEC	SDTM_LBORRESU	SDTM_LBSTRESU	SDTM_CONV_FAC	SDTM_PRECISION	SDTM_DECIMALS	SI_LBSTRESU	SI_COV_FAC
SERUM	CHEM	GGT	Gamma Glutamyl Transferase	U/L	U/L	GGT	Gamma Glutamyl Transferase	CHEMISTRY	SERUM	U/L	U/L				ukat/L	0.0167
SERUM	CHEMISTRY	GGT	Gamma Glutamyl Transferase	U/L	U/L	GGT	Gamma Glutamyl Transferase	CHEMISTRY	SERUM	U/L	U/L				ukat/L	0.0167
	CHEMISTRY	GGT	Gamma Glutamyl Transferase	U/L	U/L	GGT	Gamma Glutamyl Transferase	CHEMISTRY	SERUM	U/L	ukat/L	0.0167	0.0001	4	ukat/L	1
Serum	CHEMISTRY PANEL	RCT3	GGT	U/L	U/L	GGT	Gamma Glutamyl Transferase	CHEMISTRY	SERUM	U/L	ukat/L	0.0167	0.0001	4	ukat/L	1
	CHEMISTRY PANEL	RCT3	GGT	U/L	U/L	GGT	Gamma Glutamyl Transferase	CHEMISTRY	SERUM	U/L	ukat/L	0.0167	0.0001	4	ukat/L	1

Table 2 – Lab Standardization Effort

It is essential to harmonize all the key data points impacting on the primary and secondary objectives of the overall analysis. The look-up table approach allows the sponsor to build a robust and consistence integrated pool of data so it can be easily and correctly analyzed.

Dataset splitting consistency:

The Dataset size can be an issue to load into the eCTD tool for the submission – the size limit is 5GB.

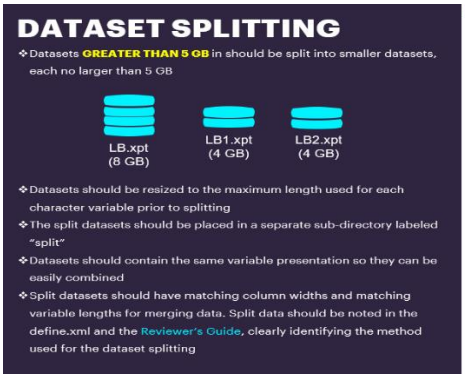
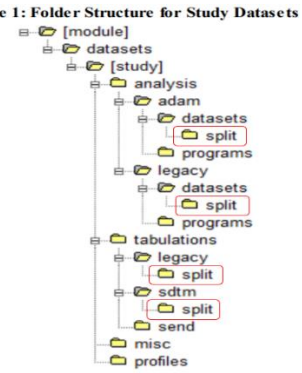


Figure 6 – Dataset Splitting guideline

Figure 1: Folder Structure for Study Datasets



If it goes beyond 5GB it must be split into smaller datasets using the appropriate data categories, subcategories, and must be documented in the define.xml and SDRG[1]. Option 3 allows a sponsor to implement this procedure with ease and in a cleaner format as the splitting happens on the standardized data. Further guidance is provided in Figure 6.

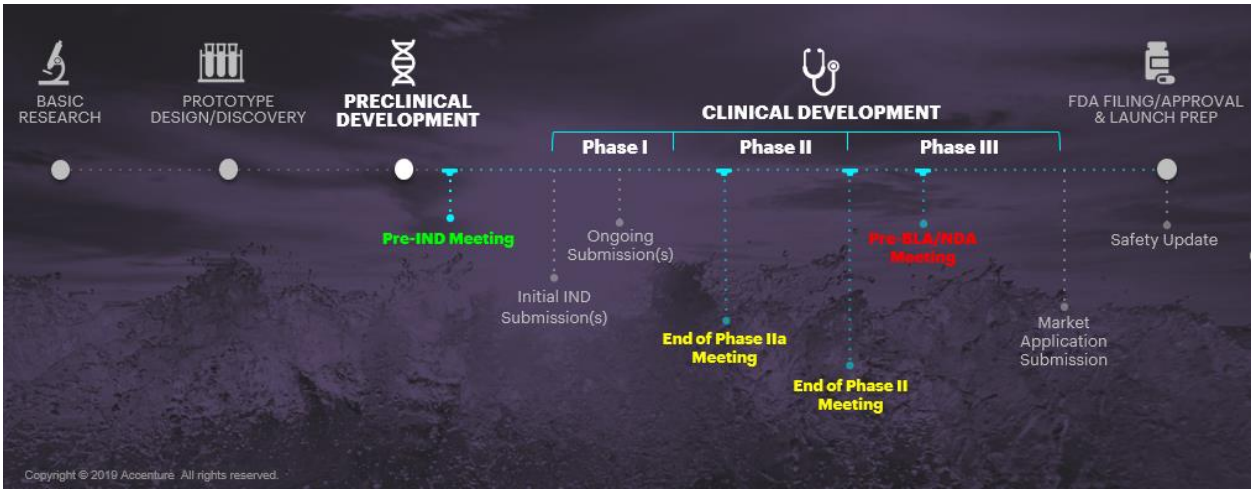
Related data publishing documents:

All the details discussed above need to be documented at someplace to sketch the flow of different transformations applied at the integrated SDTM level. The expectation is to submit the define.xml and the SDRG document along with tabulation data to complete the submission application. The define.xml details the metadata of the submitted electronic datasets and is used to describe the datasets, variables, possible values of variables when appropriate, and controlled terminologies and codes. The Define.xml v2.0 is expected as v1.0 is no longer accepted as of March 17, 2018. The Metadata Submission Guidelines, Study Data TCG and CDISC published Define.xml v2.0 release package - all provide valuable tools for proper creation. The SDRG assists FDA reviewers in conjunction with study data and the define.xml to assess the quality of your full submission package. It details planned data standards, formats, terminologies, dictionaries, and their versions. The document can be used to proactively communicate data anomalies, oddities, and transformation efforts implemented at the pooled SDTM datasets.

In the end, the most influential element in the regulatory review process is to understand the origin of the data that supports your integrated ADaM and TLFs. Traceability shows understanding of the data flow between the analysis results, analysis datasets, tabulation datasets, and the source data [1].

INTERACT WITH FDA

The sponsor should proactively stage the communication plan with FDA. Normally there 4 key meetings that occur at different milestones of the clinical development process. These are the best place to share and discuss the detailed submission plan to get agreement from FDA. The study data standardization plan is a recommended way to bring the data related topics on the table. The plan should be discussed along with quality parameters and with any potential data standardization issues at risk to get FDA reviewer's timely feedback. Ideally, the discussion on the planned standards to be used with the reviewer should occur no later than the end of phase 2. A pre-NDA/BLA meeting is considered too late to initiate the data standardization and analysis discussion.



CONCLUSION

Integrated analysis planning is a complex process by nature. The data analysis cannot be simply executed by stacking multiple tabulations or analysis data of the different trials together. Several factors must be considered and planned to preserve the integrity and accuracy of the data that are going to be analyzed. A key submission rejection criterion from FDA is the failure to define the traceability of the data from its collection through clinical study reporting, especially in cases that include the legacy trials to support the new drug and biologics submission applications. It requires step by step careful planning by understanding FDA requirements. Communication is key in this case. The data standards and the integration planning details should be discussed at the key meetings to get buy-in from FDA promptly. Option 3 allows the sponsor to implement the most logical data flow that feeds into the integration analysis, which helps the reviewer to make decision scientifically, and not by the data quality. Using the right approach from the beginning by understanding the FDA submission requirements and implementation of those during the planning can mitigate lengthy review cycles or refuse to file outcomes from FDA.

REFERENCES

1. U.S. Food & Drug Administration. Study Data Technical Conformance Guide v4.3. [Online] 03 2019. https://www.fda.gov/media/122913/download.
2. U.S. Food & Drug Administration. Data Standards Catalog v6.1 - Supported and Required Standards. [Online] 09 09, 2019. https://www.fda.gov/media/85137/download.
3. U.S. Food & Drug Administration. Janus. [Online] FDA, 11 01, 2018. https://www.fda.gov/industry/study-data-standards-resources/janus.
4. U.S. Food & Drug Administration. Study Data Standards. [Online] FDA, 06 20, 2019. https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources.

RECOMMENDED READING

1. PhUSE Wiki - https://www.phusewiki.org/wiki/index.php?title=PHUSE_Wiki
2. PhUSE efforts on the Study Data Standardization Plan - http://www.phusewiki.org/wiki/index.php?title=Study_Data_Standardization_Plan_(SDSP)
3. PharmaSUG 2017 - Paper SS11 - Documenting Traceability for the FDA: Clarifying the Legacy Data Conversion Plan & Introducing the Study Data Traceability Guide – By David C. Izard, Kristin C. Kelly, Jane A. Lozano https://www.lexjansen.com/pharmasug/2017/SS/PharmaSUG-2017-SS11.pdf

ACKNOLEDGEMENTS
The author would like to acknowledge the management at Accenture Life Sciences Data Operations group for their continuous support.

CONTACT INFORMATION
Your comments and questions are valued and encouraged. Contact the author at:

Priyank H Patel
Accenture Life Sciences
1160 W. Swedesford Road
Berwyn, PA 19312
priyank.h.patel@accenture.com