Better Health, Brighter Future

Considerations on Legacy Data Conversion to CDISC Standards for e-Data Submission for NDA in Japan

PhUSE SDE 2014
Tokyo, Japan

Takeda Pharmaceutical Company Limited
Yohei Takanami

OUTLINE

- Background
  - Current status of e-Data submission in Japan
  - Impact of e-Data submission
- Approaches for e-Data submission
  - Linear flow and Non-Linear flow
- Considerations on Legacy Data Conversion
  - Issues on materials and each Data Conversion flow
- Summary

OUTLINE

- Background
  - Current status of e-Data submission in Japan
  - Impact of e-Data submission
- Approaches for e-Data submission
  - Linear flow and Non-Linear flow
- Considerations on Legacy Data Conversion
  - Issues on materials and each Data Conversion flow
- Summary

Current Status in Japan

- Pharmaceuticals and Medical Devices Agency (PMDA) published the basic notification for e-Data submission for NDA in Japan. The e-Data submission in accordance with CDISC format to PMDA will be required by the end of FY 2016 (with a 2-year transitional period). http://www.pmda.go.jp/operations/shonin/info/iyaku/jisedai.html

- In this FY (2014), PMDA will publish additional notifications and guidance on e-Data submission.

- Pilot Project of trial submission of e-Data (2014/4 -) is currently underway. PMDA requests pharmaceutical companies to submit data in CDISC format (SDTM/ADaM datasets and Define.xml with useful documents for reviewers).

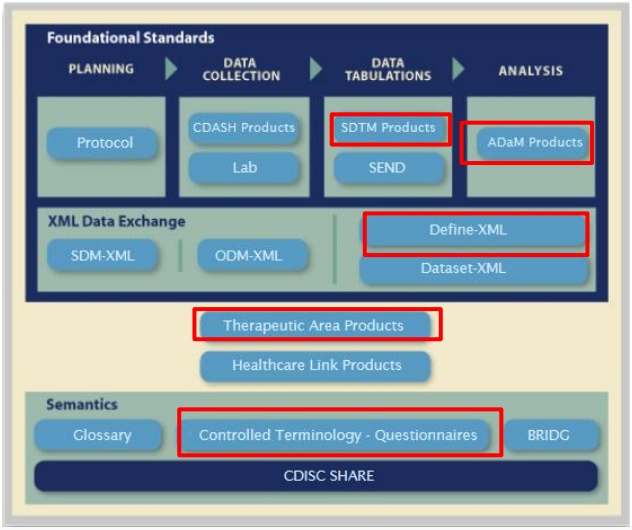
Requirements for e-Data submission in Japan

- Subject Data for submission (PMDA Basic Notification)
  - Data on results from all phase II and phase III studies (including long-term studies) that are generally regarded to be the major evidence for evaluation of efficacy, safety, and dosage and administration.
  - For study results from phase I studies and clinical pharmacology studies, results from studies listed below are required to be electronically submitted.
    - Phase I studies of oncology drugs.
    - Phase I studies that have been conducted on both Japanese and non-Japanese subjects (e.g.; in case of a strategy of global clinical trials and bridging studies).
    - QT/QTc studies based on ICH E14 guideline.

Requirements for e-Data submission in Japan

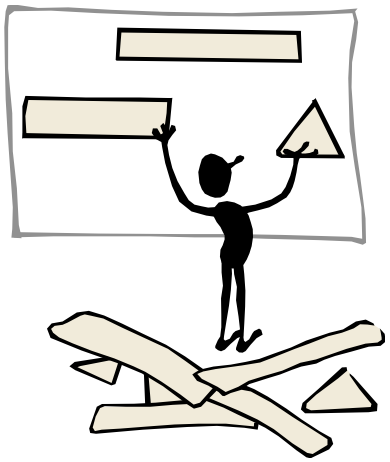
- Data standards for submission (PMDA Basic Notification)
  - Conform to CDISC standards
  - Individual study data should be prepared using the SDTM and be submitted along with the definition file for variables (e.g. Define.XML).
  - For analysis datasets, the dataset based on the ADaM should be submitted along with its definition file (e.g. Define.XML) and the program for creating the ADaM dataset.
  - For electronic data on ISS and ISE, if PMDA judges it to be necessary, in principle the dataset based on ADaM, its definition file (e.g. Define.XML), and the program for creating the ADaM dataset should be submitted
    - It is recommended that before making a submission the applicants individually consult with PMDA regarding details of what to submit.

Detail requirements will be described in the next notification and technical guide



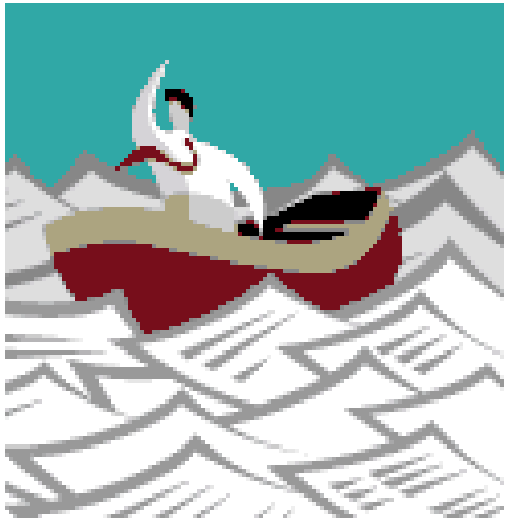
Impact of e-Data submission on PMDA and pharmaceutical companies

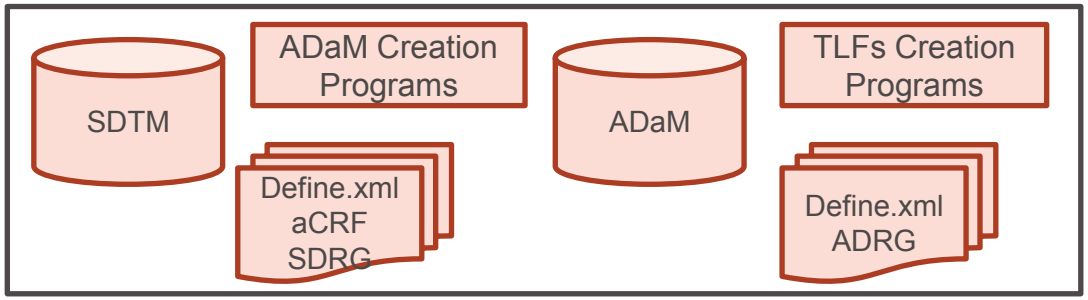
- Many changes on NDA submission and review process in Japan
  - Issues on resources and timeline for preparing and validating e-Data in accordance with CDISC standards
  - Pharmaceutical companies have to prepare many "additional" materials for e-Data submission
  - PMDA performs analyses using e-Data submitted by Pharmaceutical companies
  - Usual submission and review process are conducted in parallel
  - New PMDA Consultation Meeting (and other communication channels?) for e-Data submission
  - CDISC experts development
  - etc.



Impact of e-Data submission on pharmaceutical companies

- Preparation of materials submitted to PMDA
  - For each study
    - **SDTM datasets**
    - **Define.xml (SDTM)**
    - **Study Data Reviewer's Guide (SDRG)**
    - **Annotated CRF (aCRF)**
    - **ADaM datasets**
    - **Define.xml (ADaM)**
    - **Analysis Data Reviewer's Guide (ADRG)**
    - **ADaM Creation programs**
    - **TLFs Creation programs**
  - For ISS/ISE
    - **ADaM datasets**
    - **Define.xml**
    - **ADRG**
    - **ADaM Creation programs**
    - **TLFs Creation programs**



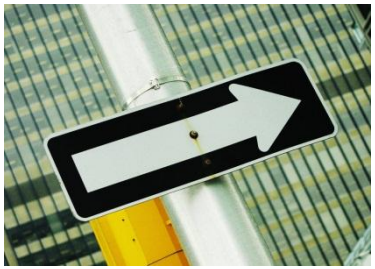


OUTLINE

- Background
  - Current status of e-Data submission in Japan
  - Impact of e-Data submission
- Approaches for e-Data submission
  - Linear flow and Non-Linear flow
- Considerations on Legacy Data Conversion
  - Issues on materials and each Data Conversion flow
- Summary

Ideal and Reality on preparing CDISC-compliant e-Data

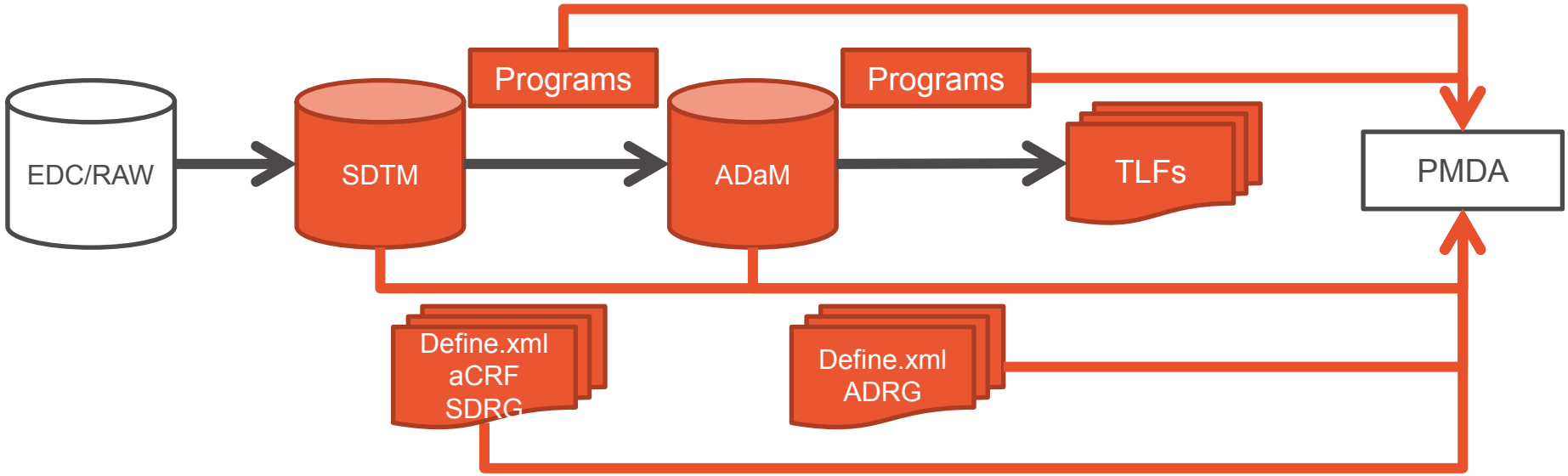
- "Linear" flow (Ideal) and "Non-Linear" flow (Reality)
  - **Linear flow**
    - Ideal flow to create CDISC-compliant e-Data
      - EDC => SDTM => ADaM => TLFs
      - No need to convert the study data for e-Data submission
      - All datasets and materials for CDSIC standards can be prepared
      - Traceability is supportable/helpful for the reviewers



  - **Non-Linear flow**
    - Need to convert the study data to the CDISC-compliant data for e-Data submission (Legacy Data Conversion)
      - **"Old" studies must be included in a submission at the beginning phase of e-Data submission in Japan**
      - **Datasets for ISS/ISE**
    - Materials in submission package should be discussed with PMDA
      - e.g. Datasets, Define.xml, Reviewers Guide

Linear flow

- Linear flow
  - Ideal flow to create CDISC-compliant e-Data
    - EDC => SDTM => ADaM => TLFs
    - No need to convert the study data for e-Data submission
    - All datasets and materials for CDSIC standards can be prepared
    - Traceability is supportable/helpful for the reviewers

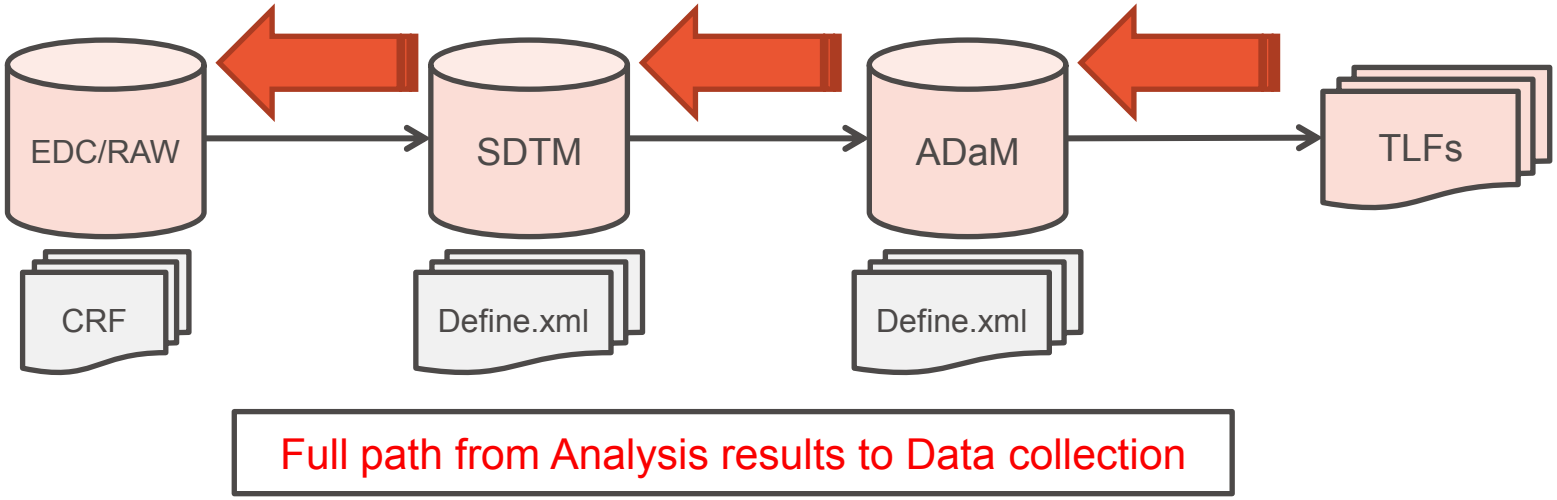


Non-Linear flow

- **Non-Linear flow with Legacy Data Conversion**
  - Conversion of Sponsor-specific datasets to CDISC-compliant data for NDA
  - **Issues well known on Legacy Data Conversion**
    - CT mapping
      - Legacy data cannot be converted to SDTM and ADaM datasets easily because CDISC has specific Controlled Terminology (CT).
    - Japanese language issue (other than English)
      - Legacy data might include Japanese characters
    - Datasets for ISS and ISE
      - Various patterns/types of source data
      - Version control (e.g. CDISC, MedDRA)
    - Traceability
      - From the ADaM standard point of view, maintaining traceability on the legacy data conversion is also a challenging issue.
    - Validation
      - Difficult to perform the complete validation for legacy data conversion (e.g. OpenCDISC Validator)

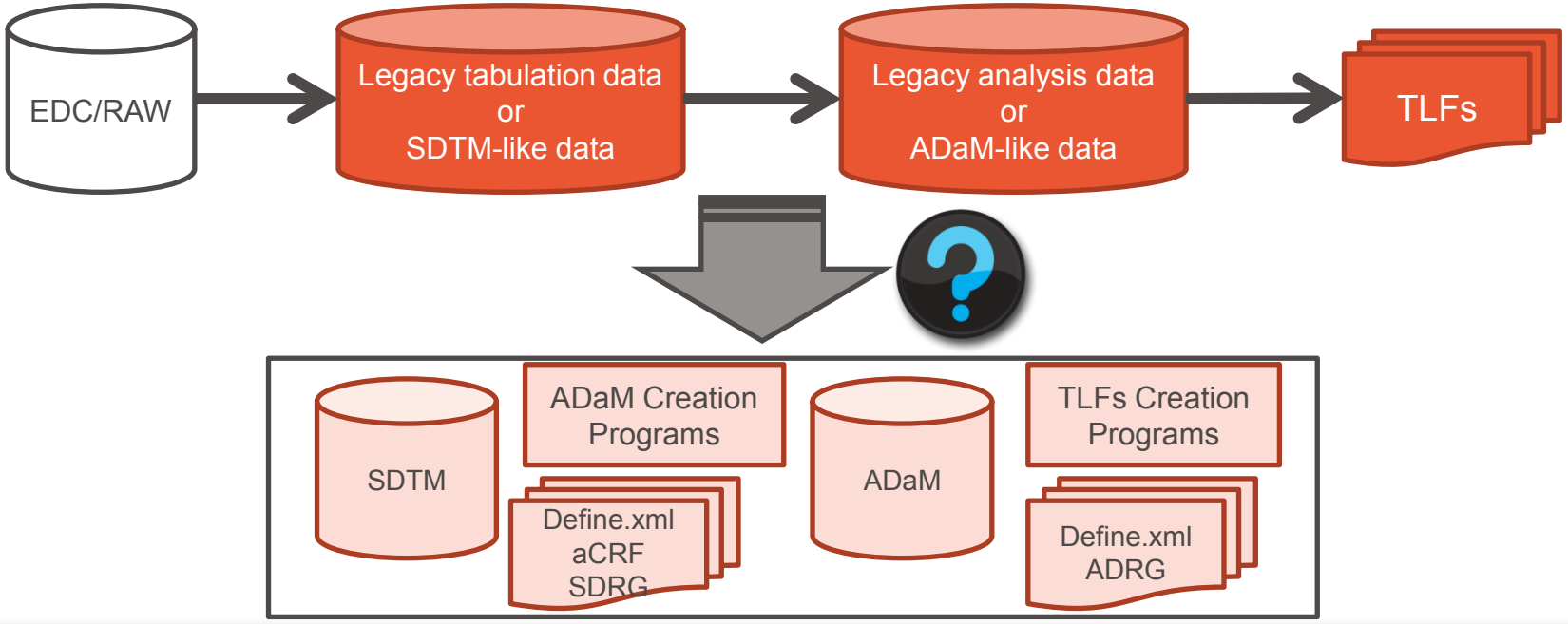
Traceability from Analysis Results

- Traceability (ADaM V2.1)
  - The concept of traceability is a cornerstone of the Analysis Data Model. This property enables the understanding of the data's lineage or the relationship between an element and its predecessor(s). Traceability facilitates transparency, which is an essential component in building confidence in a result or conclusion. Ultimately, traceability in ADaM permits the understanding of the relationship among the analysis results, the analysis datasets, and the SDTM domains.



Non-Linear flow

- Non-Linear flow
  - Approaches for Legacy Data Conversion
    - SDTM-like model (Conversion from SDTM-like to SDTM)
    - Sequential conversion of SDTM and ADaM standards
    - Independent conversion of SDTM and ADaM standards
    - Creation of ISS/ISE ADaM datasets from different types source data
  - **How/What do sponsors prepare the submission package?**



OUTLINE

- Background
  - Current status of e-Data submission in Japan
  - Impact of e-Data submission
- Approaches for e-Data submission
  - Linear flow and Non-Linear flow
- Considerations on Legacy Data Conversion
  - Issues on materials and each Data Conversion flow
- Summary

Considerations on e-Data submission package

- Patterns of Data flow

Pattern	Data flow
Linear	1. Raw=>SDTM=>ADaM=>TLFs
Non-Linear	2. SDTM-like model (Conversion from SDTM-like to SDTM)
	3. Sequential Legacy Data Conversion of SDTM and ADaM standards
	4. Independent conversion of SDTM and ADaM standards
	5. Creation of ISS/ISE ADaM datasets from different types source data

- Materials for e-Data submission

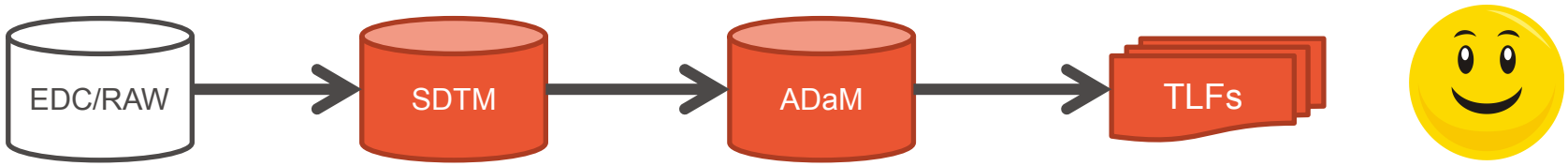
Data Model	Materials
SDTM	Datasets
	Define.xml
	Annotated CRF
	Study Data Reviewer's Guide
ADaM	Datasets
	Define.xml
	ADaM Creation Programs
	TLFs Creation Programs
	Analysis Data Reviewer's Guide

Considerations on Legacy Data Conversion

- Linear
    1. Raw=>SDTM=>ADaM=>TLFs
- Non-Linear
    2. SDTM-like model (Conversion from SDTM-like to SDTM)
    3. Sequential Legacy Data Conversion of SDTM and ADaM standards
    4. Independent conversion of SDTM and ADaM standards
    5. Creation of ISS/ISE ADaM datasets from different types source data

Linear flow

- Linear flow
  - Sponsor can prepare all the materials with CDISC-compliant data



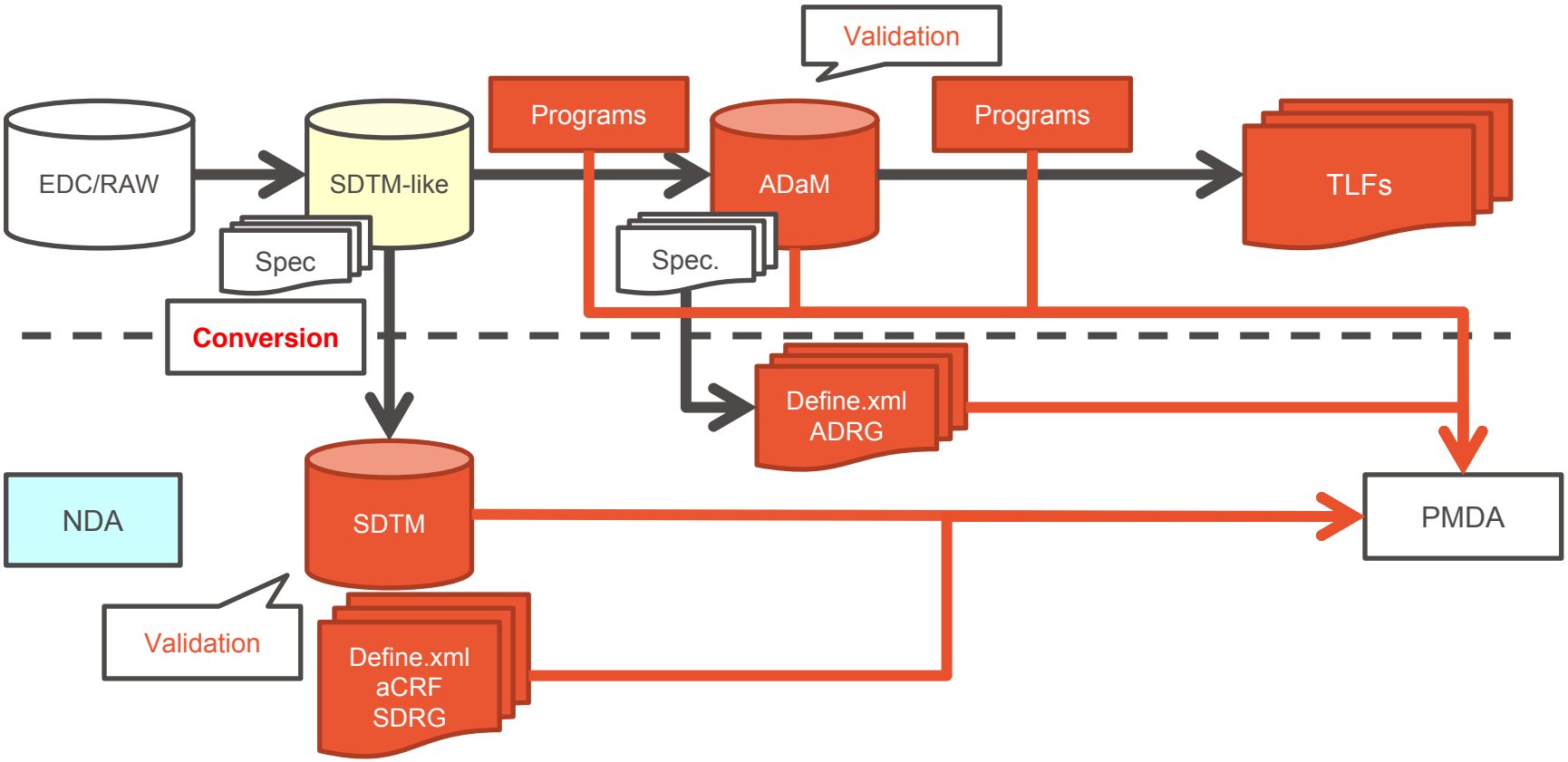
Data Model	Materials	Status	Note
SDTM	Datasets	OK	
	Define.xml	OK	
	Annotated CRF	OK	
	Study Data Reviewer's Guide	OK	
ADaM	Datasets	OK	
	Define.xml	OK	
	ADaM Creation Programs	OK	
	TLFs Creation Programs	OK	
	Analysis Data Reviewer's Guide	OK	

Considerations on Legacy Data Conversion

- Linear
  1. Raw=>SDTM=>ADaM=>TLFs
- Non-Linear
  2. SDTM-like model (Conversion from SDTM-like to SDTM)
  3. Sequential Legacy Data Conversion of SDTM and ADaM standards
  4. Independent conversion of SDTM and ADaM standards
  5. Creation of ISS/ISE ADaM datasets from different types source data

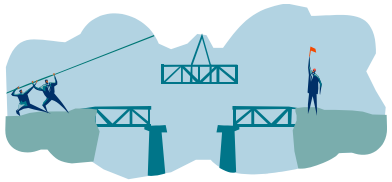
Non-Linear flow 1

- Convert the SDTM-like to pure SDTM
  - Submit the ADaM datasets without conversion to reduce the workload



Non-Linear flow 1

- Convert the SDTM-like to pure SDTM
  - Traceability issues between ADaM and SDTM
  - ADaM Define.xml describes "pure" SDTM datasets and variables for reviewers to keep the traceability as much as possible
  - ADaM creation programs are based on SDTM-like datasets
  - Validation of ADaM datasets should be done with pure SDTM?



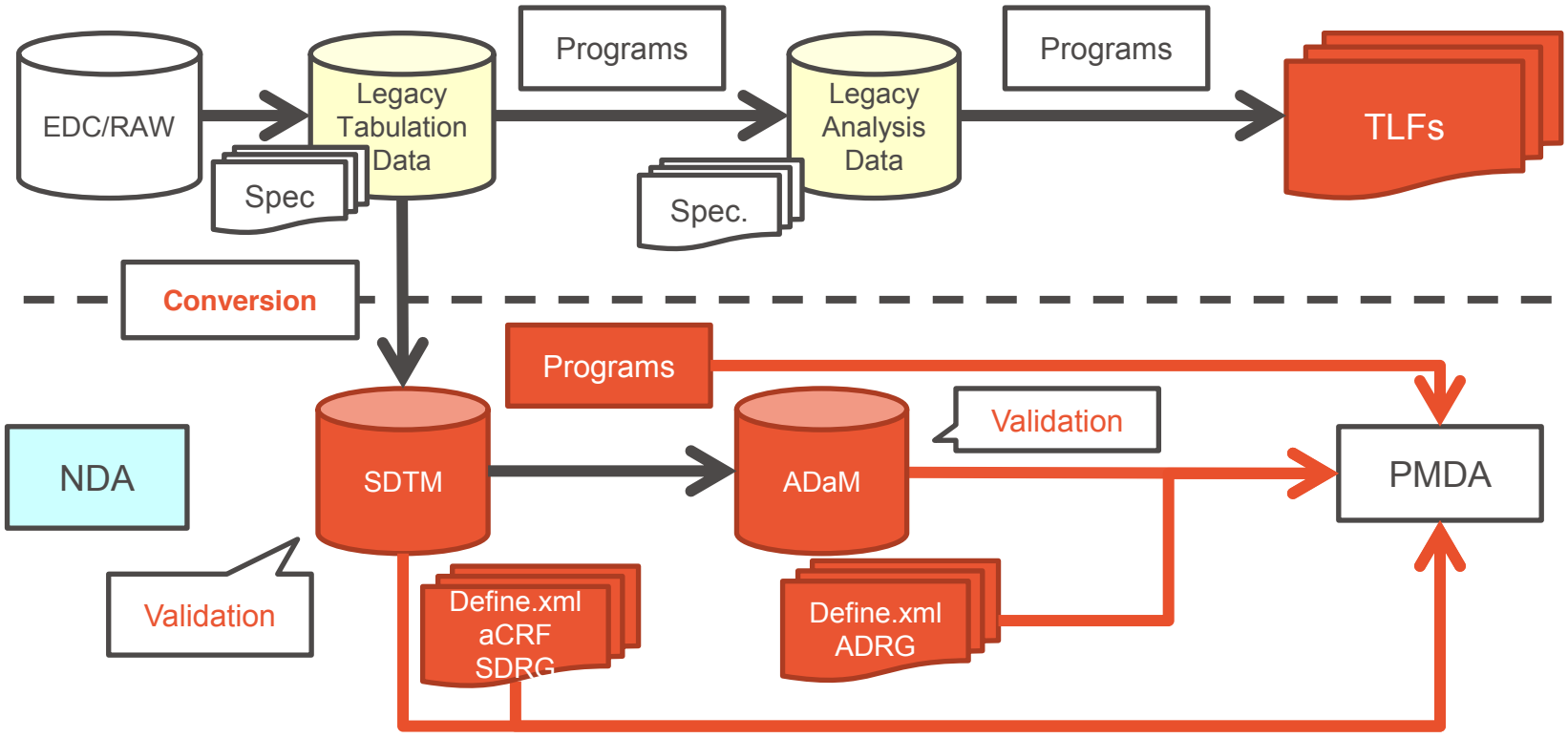
Data Model	Materials	Status	Note
SDTM	Datasets	OK	Converted from SDTM-like datasets
	Define.xml	OK	For pure SDTM
	Annotated CRF	OK	For pure SDTM
	Study Data Reviewer's Guide	OK	For pure SDTM
ADaM	Datasets	Cond	Create from SDTM-like (traceability is not perfect)
	Define.xml	Cond	Source should be pure SDTM (traceability is not perfect)
	ADaM Creation Programs	Cond	From SDTM-like datasets (traceability is not perfect)
	TLFs Creation Programs	OK	
	Analysis Data Reviewer's Guide	Cond	Pure SDTM should be described as the source data

Considerations on Legacy Data Conversion

- Linear
  1. Raw=>SDTM=>ADaM=>TLFs
- Non-Linear
  2. SDTM-like model (Conversion from SDTM-like to SDTM)
  3. Sequential Legacy Data Conversion of SDTM and ADaM standards
  4. Independent conversion of SDTM and ADaM standards
  5. Creation of ISS/ISE ADaM datasets from different types source data

Non-Linear flow 2

- Legacy Data to SDTM and SDTM to ADaM
  - Convert the Legacy Data to SDTM and create ADaM from the SDTM



Non-Linear flow 2

- Legacy Data to SDTM and SDTM to ADaM
  - CT issues for the conversion of Legacy Data to SDTM
  - Language issue
  - No TLFs creation programs

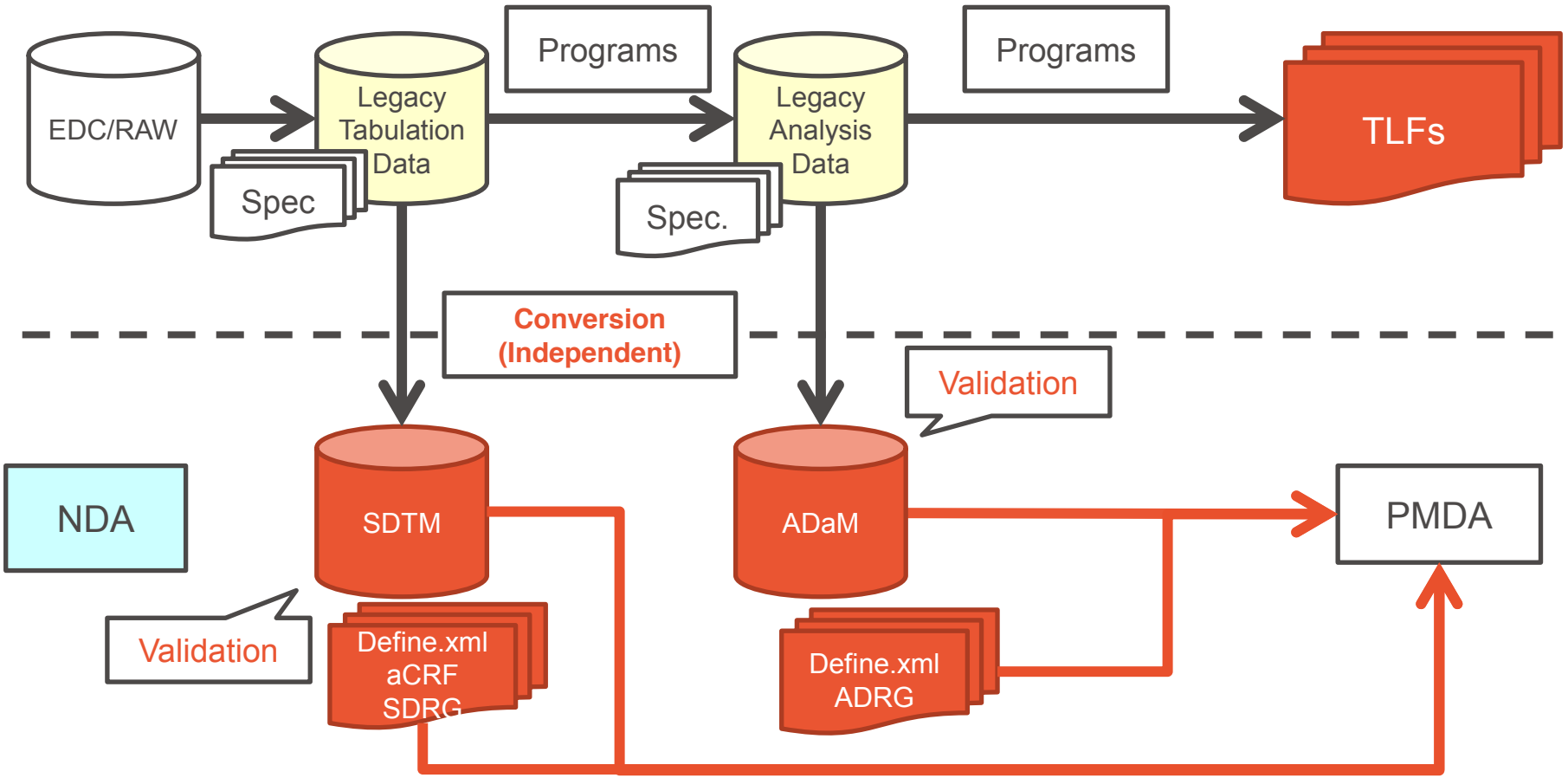
Data Model	Materials	Status	Note
SDTM	Datasets	Cond	Legacy Data Conversion
	Define.xml	Cond	Legacy Data Conversion
	Annotated CRF	Cond	Legacy Data Conversion
	Study Data Reviewer's Guide	Cond	Legacy Data Conversion
ADaM	Datasets	Cond	From SDTM
	Define.xml	Cond	From SDTM
	ADaM Creation Programs	Cond	From SDTM
	TLFs Creation Programs	None	Using legacy data (no traceability to ADaM)
	Analysis Data Reviewer's Guide	Cond	SDTM is described as source data

Considerations on Legacy Data Conversion

- Linear
  1. Raw=>SDTM=>ADaM=>TLFs
- Non-Linear
  2. SDTM-like model (Conversion from SDTM-like to SDTM)
  3. Sequential Legacy Data Conversion of SDTM and ADaM standards
  4. Independent conversion of SDTM and ADaM standards
  5. Creation of ISS/ISE ADaM datasets from different types source data

Non-Linear flow 3

- Legacy Tabulation Data to SDTM and Legacy Analysis Data to ADaM
  - Convert the Legacy Data to SDTM and Legacy Analysis Data to ADaM **independently**



Non-Linear flow 3

- Legacy Tabulation Data to SDTM and Legacy Analysis Data to ADaM
  - No traceability between ADaM and SDTM
  - What should be described in the ADaM Define.xml and ADRG?
  - No ADaM creation programs from SDTM
  - Japanese language should be converted in parallel?
  - ADaM Validation cannot be perfect

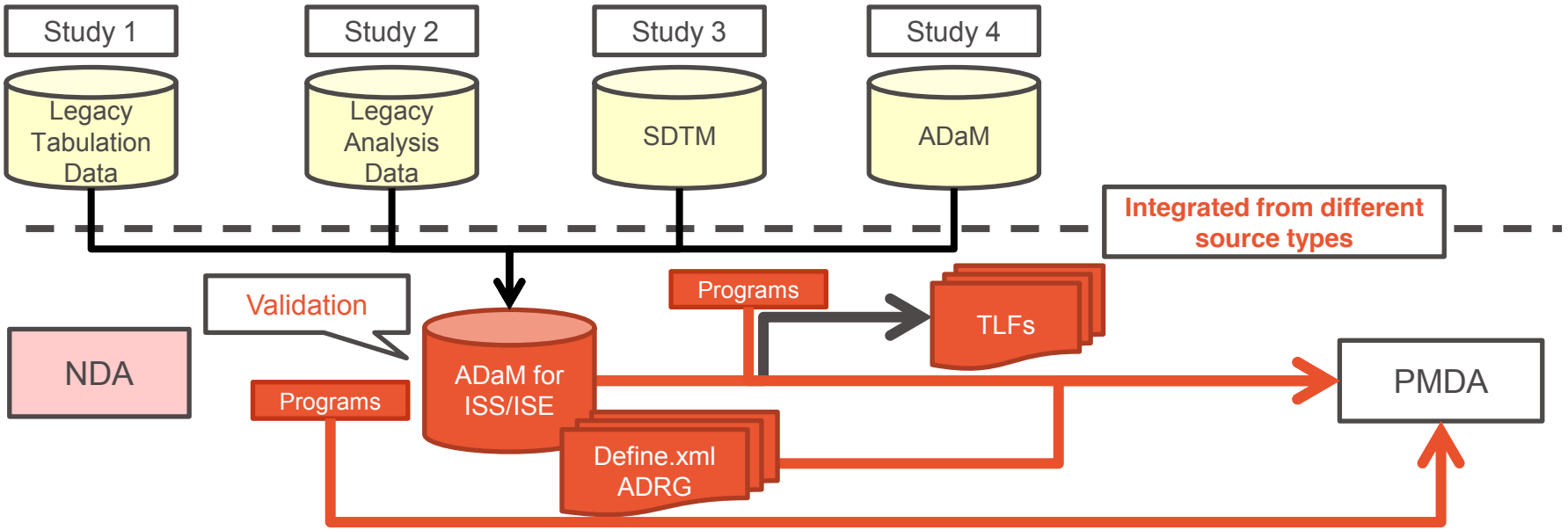
Data Model	Materials	Status	Note
SDTM	Datasets	Cond	Legacy Data Conversion
	Define.xml	Cond	Legacy Data Conversion
	Annotated CRF	Cond	Legacy Data Conversion
	Study Data Reviewer's Guide	Cond	Legacy Data Conversion
ADaM	Datasets	Cond	Legacy Data Conversion (no traceability)
	Define.xml	Cond	Difficult to describe the source data (no traceability)
	ADaM Creation Programs	None	Using legacy data (no traceability to SDTM)
	TLFs Creation Programs	None	Using legacy data (no traceability to ADaM)
	Analysis Data Reviewer's Guide	Cond	Difficult to create (no traceability to SDTM)

Considerations on Legacy Data Conversion

- Linear
  1. Raw=>SDTM=>ADaM=>TLFs
- Non-Linear
  2. SDTM-like model (Conversion from SDTM-like to SDTM)
  3. Sequential Legacy Data Conversion of SDTM and ADaM standards
  4. Independent conversion of SDTM and ADaM standards
  5. Creation of ISS/ISE ADaM datasets from different types source data

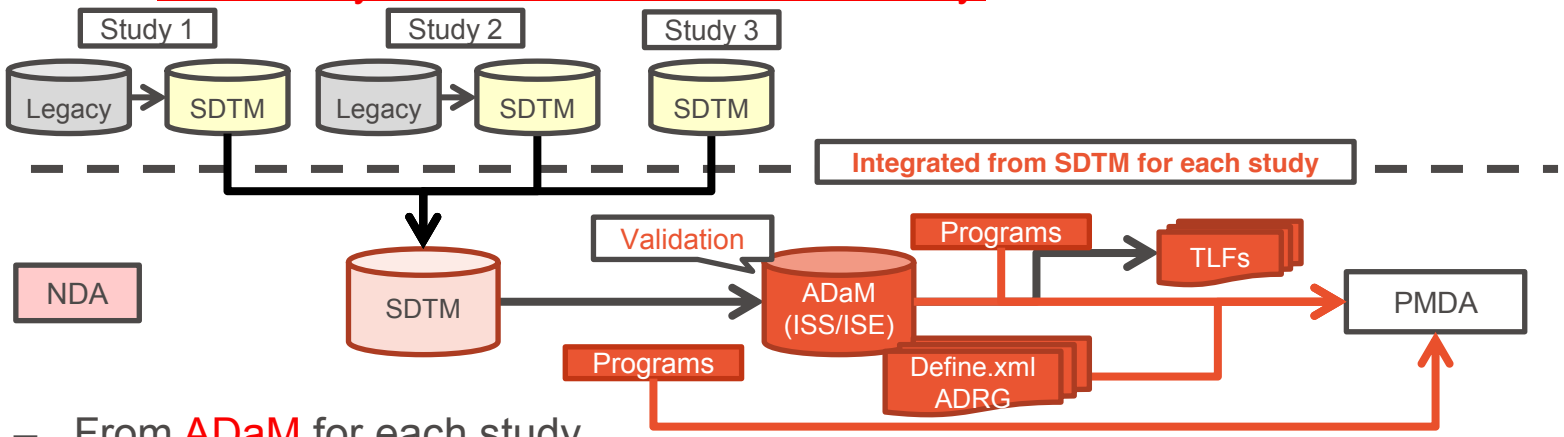
Non-Linear flow 4

- ISS and ISE
  - Many issues on ISS/ISE ADaM datasets, documents, programs and validation
  - What should be the source data?
    - Controversial issue for the creation of ISS/ISE ADaM datasets
  - From different type of source data
    - Traceability issues and complex programs for ADaM creation

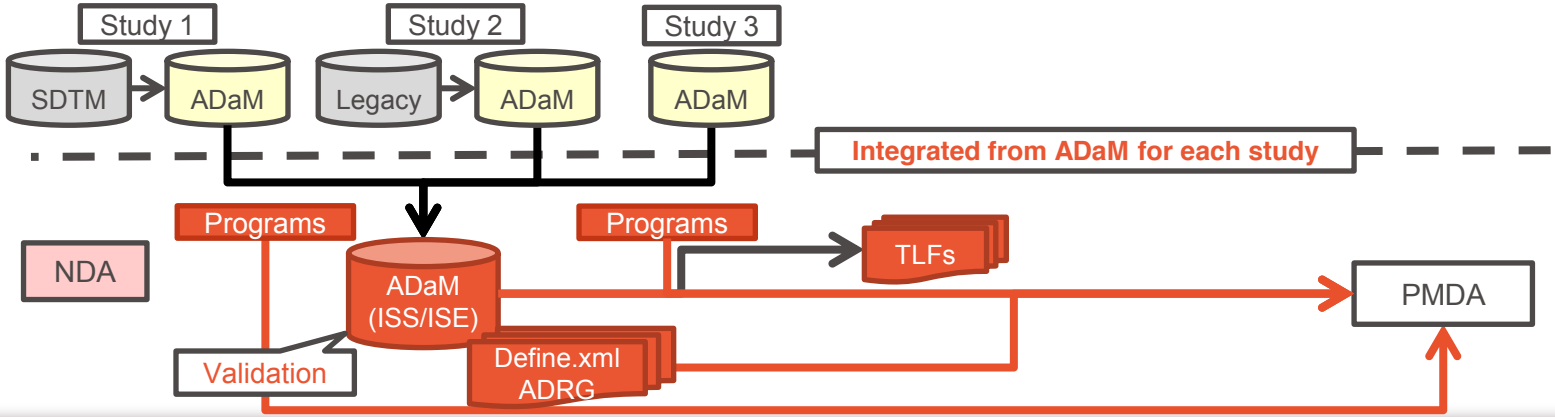


Non-Linear flow 4

- Source Data: SDTM or ADaM?
  - From integrated SDTM (from SDTM for each study)
    - Traceability is better, but Workload is heavy.



  - From ADaM for each study
    - Traceability is kept between ADaM datasets. Workload is not so heavy.



Non-Linear flow 4

- ISS and ISE
  - Both Legacy and CDISC-compliant data might be included in the source data.
    - Type of Source data is controvertible (integrated SDTM or ADaM for each study or sponsor specific legacy data)
    - Traceability issues
  - What should be described in the ADaM Define.xml and ADRG?
    - Define.xml is one file, so it is difficult to describe the details of source/derivation column
  - ADaM creation programs integrate various types of source data
  - ADaM Validation cannot be perfect (e.g. OpenCDISC validator)
  - Some PMDA notifications require the Japanese output in CTD
    - Output is Japanese, but Japanese character should not be included in the dataset
  - Version control (e.g. SDTM, ADaM, CT, MedDRA)

Data Model	Materials	Status	Note
ADaM	Datasets	Cond	Legacy Data Conversion might be included (data type might differ by each study, less traceability)
	Define.xml	Cond	Difficult to describe the source data (data type might differ by each study)
	ADaM Creation Programs	Cond	Source data might not be SDTM
	TLFs Creation Programs	OK	If ISS/ISE is performed using the integrated ADaM datasets
	Analysis Data Reviewer's Guide	Cond	Difficult to create (source data might not be SDTM)

OUTLINE

- Background
  - Current status of e-Data submission in Japan
  - Impact of e-Data submission
- Approaches for e-Data submission
  - Linear flow and Non-Linear flow
- Considerations on Legacy Data Conversion
  - Issues on materials and each Data Conversion flow
- Summary

Summary

- Impact of e-Data submission on NDA in Japan
  - Workload increases in PMDA and Pharmaceutical companies
  - Many changes on NDA submission and review process in Japan

- Considerations on Legacy Data Conversion
  - One of the challenging approach for preparing the CDISC-compliant e-Data
    - Need to convert the legacy data to create the CDISC-compliant e-Data especially at the beginning phase of e-Data submission in Japan
  - Several patterns and controversial issues on legacy data conversion including ISS/ISE ADaM datasets creation
    - e.g. Traceability, Validation, Documentation, Programs
  - Need to discuss the e-Data submission package with PMDA
    - New consultation meeting
    - Other communication channels (e.g. email, teleconference)?

References

- PMDA HP for e-Data submission
  http://www.pmda.go.jp/operations/shonin/info/iyaku/jisedai.html
- FDA (2014). STUDY DATA TECHNICAL CONFORMANCE GUIDE (DRAFT)
  http://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM384744.pdf
- CDISC HP
  http://www.cdisc.org/
- PhUSE White Paper. Traceability and Data Flow Best Practices for Basic Linear Data Flow
  http://www.phusewiki.org/wiki/images/4/49/Basic_Linear_Data_Flow_White_Paper_2014-10-01.pdf
- PhUSE White Paper. Study Level Traceability Considerations Best Practices for Non-Linear Data Flow
  http://www.phusewiki.org/wiki/images/b/b3/Study_Level_Traceability_White_Paper_August_2014.pdf
- Legacy Data Conversion Plan & Report (PhUSE Wiki)
  http://www.phusewiki.org/wiki/index.php?title=Legacy_Data_Conversion_Plan_%26_Report
- Joerg Guettner, Alexandru Cuza. PhUSE 2012. ADaM or SDTM? A Comparison of Pooling Strategies for Integrated Analyses in the Age of CDISC
  http://www.phusewiki.org/docs/2012/PAPERS/CD/CD01.pdf
- Chris Holland, Jack Shostak (2012). Implementing CDISC Using SAS: An End-to-End Guide. Sas Inst

End of file

Takeda Pharmaceutical Company Limited