Paper PP18

PHUSE CDISC Implementation Primer Project Overview and Update

Beate Hientzsch, PHUSE Education Director
Nick De Donder, Business & Decision Life Sciences, Brussels, Belgium

ABSTRACT

As the CDISC implementation community increases, new implementers, especially those not directly connected to the pharmaceutical industry and regulators, need a starting point for their implementation journey. At the 2018 PHUSE CSS, the CDISC Implementation Primer project was launched as part of the PHUSE Working Group "Optimizing the use of Data Standards". The goal of the project is to provide new CDISC implementers a primer to start their journey. This Primer (e.g. Wiki, White Paper, slide deck, quickstart, guides, etc.) will be developed to address these implementation needs and will be made freely available to the public through the PHUSE website. Topics to be addressed include: how to get started with CDISC, understanding the linking of models and implementation guides, compliance, and traceability. The CDISC Implementation Primer is anticipated to be completed in spring 2020.

INTRODUCTION

As the CDISC implementation community increases, new implementers, especially those not directly connected to the pharmaceutical industry and regulators, need a starting point. The goal of this PHUSE project is to provide new CDISC implementers a freely available primer developed to address these implementation needs.
The CDISC Implementation Primer project was launched at the 2018 PHUSE CSS as part of the PHUSE Working Group "Optimizing the use of Data Standards", https://www.phuse.eu/optimizing-data-standards.
The members of the team decided to first focus on the four topics outlined below.

HOW TO GET STARTED WITH CDISC

Data Management and Statistical Programming but also other functions play vital roles in the standardization of clinical & non clinical trial data. At the first glance the various CDISC data models and their linkage can be overwhelming. Some guidance is thus needed to help understanding which standards are relevant at the different stages in the data flow from collection to analysis, reporting and submission.
Data standardization involves the annotation of CRF as per SDTM, dataset conversion & creation using CDISC Standards (SEND, SDTM & ADaM) and creation of supportive documents like reviewer guides & metadata files (define.xml).

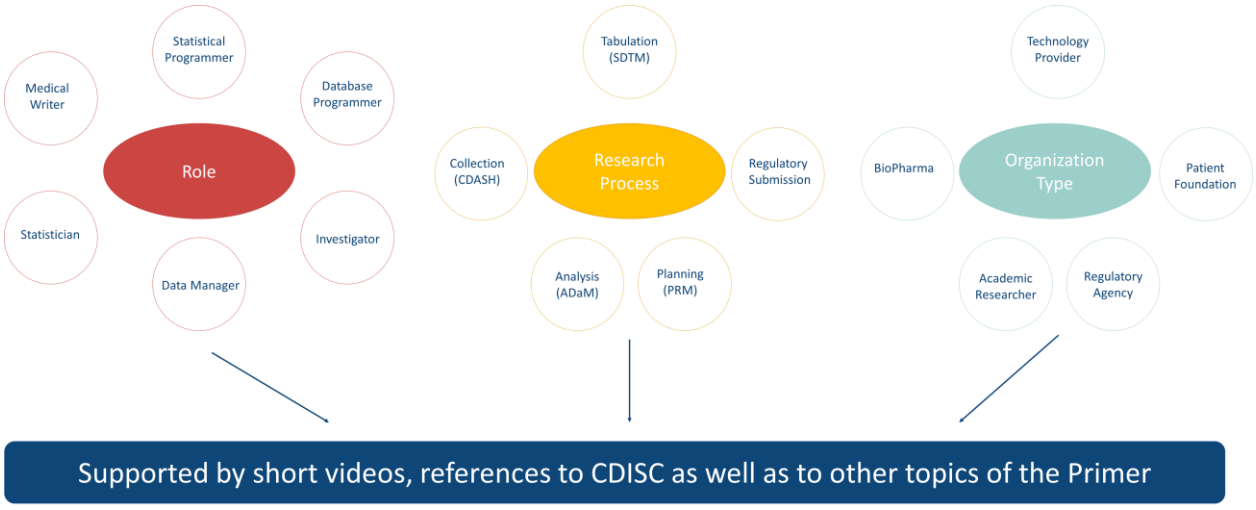

Figure 1: Understanding CDISC standards by role, clinical research process, or organization type

The activities throughout the clinical development process include but are not limited to

- Creation of annotated CRFs using SDTM standards.
- Development and validation of SDTM domains using the CDISC implementation guide & CDISC controlled terminology
- Development and validation of ADaM datasets in accordance to CDISC implementation standards
- Development and validation of Tables, Listings and Figures, eventually the creation of analysis results metadata.
- Creation and validation of Define.xml
- Creation and QC of cSDRG (clinical study data reviewer's guide) & ADRG (analysis data reviewer's guide)
- Compliance checking against regulatory requirements using different tools

COMPLIANCE

The regulatory agencies have published requirements for eSubmissions. These resources are available through their respective website. When CDISC standards are required, the compliance to those needs to be ensured and checked before submission.

This subproject of the PHUSE Implementation Primer Project is currently paused and looking for volunteers.

Figure 2: Compliance is the core piece to avoid refusal to file in the drug application process.

UNDERSTANDING THE LINK BETWEEN THE CDISC STANDARDS

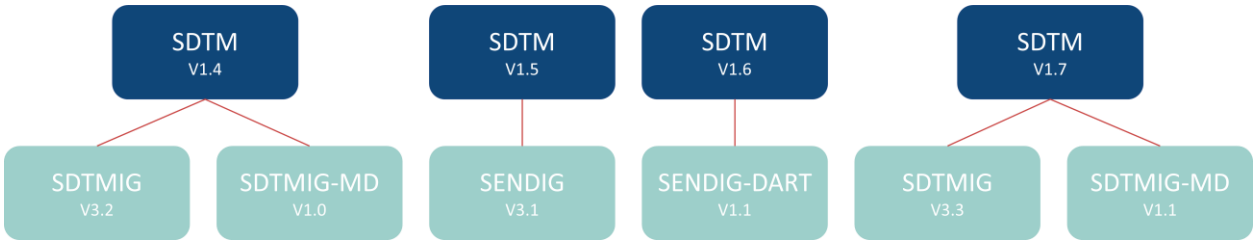

Figure 3: Link between SDTM models and Implementation Guides

There is a general misunderstanding that one SDTM (Study Data Tabulation Model) is linked to one Implementation Guide (IG). Besides supporting the SDTMIG V3.2, model V1.4 (released in 2013) was also created to support the SDTMIG-MD V1.0 (implementation guide for Medical Devices). In 2016 and 2017, two new versions of the model have been released but no associated implementation guide. These 2 models were published mainly to support the Standard for the Exchange of Non-clinical Data (SEND). V1.5 was released to support SENDIG V3.1 and V1.6 was released to support SENDIG-DART V1.1 (Developmental And Reproductive Toxicology).
In 2018 V1.7 of the SDTM was released together with SDTMIG V3.3 and SDTMIG-MD V1.1.

TRACEABILITY

Documenting the relationships from analysis results to source data as collected is key to ensure traceability on data-point and metadata level.

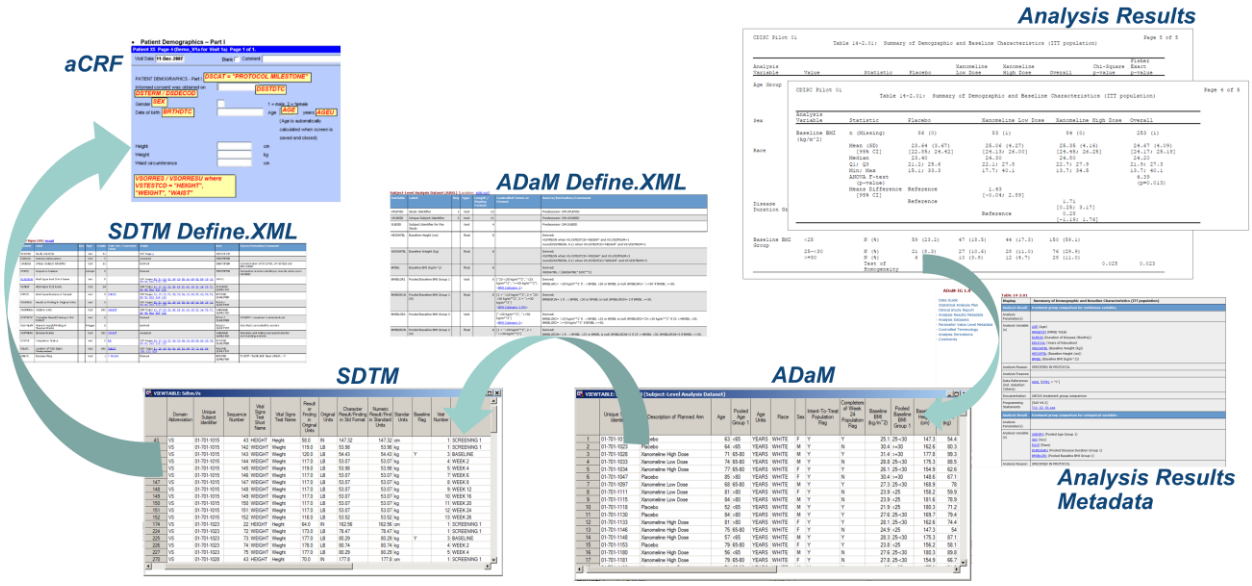

Figure 4: Traceability overview

Traceability is the property that enables the understanding where the analysis results come from. It facilitates transparency and is an essential component for building confidence in analysis results. The traceability shown in Figure 4 shows the linear data flow. Starting with the protocol, one defines the analysis need of tables, listings and figures (TFLs), and the ADaM datasets needed to support the TLFs. Based on the ADaM datasets, one can trace back the required SDTM datasets, and the corresponding case report forms and fields via the metadata described in Define-XML.
There are 2 ways of describing traceability: data point traceability and metadata traceability.

METADATA TRACEABILITY

The Metadata traceability includes relationship between an analysis result and analysis dataset(s), and relationship of the analysis variables to its source dataset(s) and variable(s) in SDTM or other ADaM dataset(s). The traceability includes relationship of the SDTM variable to collected data in CRF or to other SDTM variable(s) through derivation. This traceability is established by describing (via metadata) the algorithms used, or steps taken to derive or populate an analysis variable from its immediate predecessor.
Metadata traceability is also used to establish the relationship between an analysis result and ADaM dataset(s).

DATAPOINT TRACEABILITY

Data point traceability points directly to the specific predecessor record(s) and should be implemented if practical and feasible. This level of traceability can be very helpful when trying to trace a complex data manipulation path.
This traceability is established by providing clear links in the data (e.g., use of --SEQ variable) to the specific data values used as input for an analysis value. The BDS and OCCDS structures were designed to enable data point traceability back to predecessor data.
It may not always be practical or feasible to provide data point traceability via record identifier variables from the source dataset(s). However, metadata traceability must always clearly explain how an analysis variable is populated regardless of whether data point traceability is also provided. Very complex derivations may require the creation of intermediate analysis datasets.

CONCLUSION

The PHUSE CDISC Implementation Primer Project was initiated to help newcomers at various places to find their

way through the CDISC standards. Over the last 18 months, the team made good progress in preparing deliverables for the community. The subteam working on traceability already collected public comments.
Another update of the project status will be given at the PHUSE US Connect in 2020.

REFERENCES

https://www.phuse.eu/optimizing-data-standards
https://www.cdisc.org/

CONTACT INFORMATION

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

Beate Hientzsch
PHUSE Education Director
beate.hientzsch@phuse.eu

Nick De Donder
Business & Decision Life Sciences
Sint-Lambertusstraat 141
1200 Brussels
Belgium
+32 476 82 20 27
nick.dedonder@businessdecision.be
https://www.businessdecision-lifesciences.com/