PharmaSUG 2023 - Paper MM-327

All You Need to Know about the New CDISC Analysis Result Standards!

Bhavin Busa, Clymb Clinical
Richard Marshall, CDISC
Bess LeRoy, CDISC

ABSTRACT

The CDISC Analysis Results Standard (ARS) Team is developing standards to improve and facilitate the automation, reproducibility, reusability, and traceability of analysis results (Tables, Figures and Listings, aka TFL). The team is working towards this objective by creating a logical model to incorporate the elements for both analysis results and associated metadata.

In this paper, we will provide a comprehensive overview of the CDISC Analysis Results logical model, including its background and development. We will discuss the current state workflow and challenges faced by analysts, as well as the future state envisioned with the introduction of the machine-readable analysis results metadata and analysis results data structure to automate the generation of statistical outputs and displays. We will describe the model's elements using examples to give a better understanding of the model itself and how it can improve traceability, reproducibility, and quality in clinical trial analysis and reporting.

INTRODUCTION

Analysis results play a crucial role in the drug development process, providing essential information for regulatory submission and decision-making. However, the current state of analysis results reporting is suboptimal, with limited standardization, lack of automation, and poor traceability. Currently, analysis results (tables, figures, and listings) are often presented in static, PDF-based reports that are difficult to navigate and vary between sponsors. Moreover, these reports are expensive to generate and offer limited reusability. To address these issues, the CDISC ARS team has been working on developing standards to support consistency, traceability, and reuse of results data. In this paper, we will outline the current state of analysis results reporting, the goals of the ARS team, and propose a new workflow that will revolutionize the way analysis results are generated.

Current state

Currently, analysis results are created as static PDF documents that may contain hundreds of tables. These tables are often difficult to navigate and there is significant variability between sponsors. Generating these reports is expensive and they are typically only used once, offering limited reusability.

The current workflow of generating analysis results involves the end user generating the Analysis Data Model (ADaM) dataset, followed by generating the display in a static format such as RTF or PDF using the ADaM dataset. The Analysis Results Metadata (ARM) for define.xml is retrospectively generated, which provides high-level metadata about analysis displays and results (Figure 1); however, there is no formal model or structures to describe analysis results metadata and analysis results data which leaves a gap in standardization.

The current process is expensive, time-consuming, and lacks automation and traceability, leading to unnecessary variation in analysis results reporting.

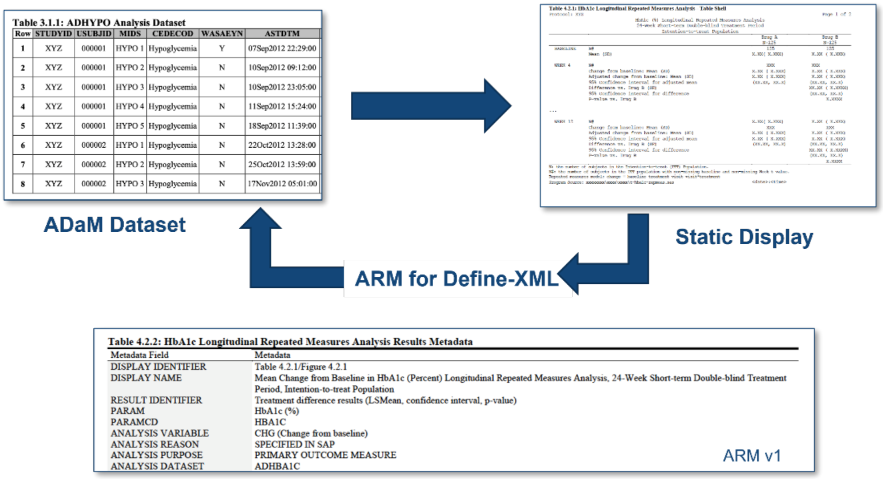


Table 3.1.1: ADHYPO Analysis Dataset

Row	STUDYID	USUBJID	MIDS	CEDECOD	WASAEYN	ASTDTM
1	XYZ	000001	HYPO 1	Hypoglycemia	Y	07Sep2012 22:29:00
2	XYZ	000001	HYPO 2	Hypoglycemia	N	10Sep2012 09:12:00
3	XYZ	000001	HYPO 3	Hypoglycemia	N	10Sep2012 23:05:00
4	XYZ	000001	HYPO 4	Hypoglycemia	N	11Sep2012 15:24:00
5	XYZ	000001	HYPO 5	Hypoglycemia	N	18Sep2012 11:39:00
6	XYZ	000002	HYPO 1	Hypoglycemia	N	22Oct2012 13:28:00
7	XYZ	000002	HYPO 2	Hypoglycemia	N	25Oct2012 13:59:00
8	XYZ	000002	HYPO 3	Hypoglycemia	N	17Nov2012 05:01:00

Table 4.2.1: HbA1c Longitudinal Repeated Measures Analysis - Table Shell

```
Protocol: XYZ                                                          Page 1 of 2
                    HbA1c (%) Longitudinal Repeated Measures Analysis
                  24-week Short-term Double-blind Treatment Period
                             Intention-to-treat Population
                                                              Drug A        Drug B
                                                              N=125         N=125
  BASELINE    N#                                               125           125
              Mean (SD)                                    X.XX( X.XXX)  X.XX ( X.XXX)

  WEEK 4      N#                                               XXX          XXX
              Change from baseline: Mean (SD)              X.XX ( X.XXX)  X.XX ( X.XXX)
              Adjusted change from baseline: Mean (SD)     X.XX ( X.XXX)  X.XX ( X.XXX)
              95% confidence interval for adjusted mean    (XX.XX, XX.X)  (XX.XX, XX.X)
              Difference vs. Drug B (SE)                                 XX.XX ( X.XXXX)
              95% confidence interval for difference                     (XX.XX, XX.X)
              P-value vs. Drug B                                             X.XXXX
...

  WEEK 12     N#                                           X.XX( X.XXX)  X.XX ( X.XXX)
              Change from baseline: Mean (SD)                  XXX           XXX
              Adjusted change from baseline: Mean (SD)     X.XX ( X.XXX)  X.XX ( X.XXX)
              95% Confidence interval for adjusted mean    X.XX ( X.XXX)  X.XX ( X.XXX)
              Difference vs. Drug B (SE)                   (XX.XX, XX.X)  (XX.XX, XX.X)
              95% Confidence interval for difference                     XX.XX ( X.XXXX)
              P-value vs. Drug B                                         (XX.XX, XX.X)
                                                                             X.XXXX
N: the number of subjects in the Intention-to-treat (ITT) Population.
N#: the number of subjects in the ITT population with non-missing baseline and non-missing Week t value.
Repeated measures model: change = baseline treatment visit visit*treatment
Program Source: [illegible]                              <date>:<time>
```

Table 4.2.2: HbA1c Longitudinal Repeated Measures Analysis Results Metadata

Metadata Field	Metadata
DISPLAY IDENTIFIER	Table 4.2.1/Figure 4.2.1
DISPLAY NAME	Mean Change from Baseline in HbA1c (Percent) Longitudinal Repeated Measures Analysis, 24-Week Short-term Double-blind Treatment Period, Intention-to-treat Population
RESULT IDENTIFIER	Treatment difference results (LSMean, confidence interval, p-value)
PARAM	HbA1c (%)
PARAMCD	HBA1C
ANALYSIS VARIABLE	CHG (Change from baseline)
ANALYSIS REASON	SPECIFIED IN SAP
ANALYSIS PURPOSE	PRIMARY OUTCOME MEASURE
ANALYSIS DATASET	ADHBA1C

Figure 1: Current State of Display and ARM Generation

Future State

Our vision for the future state of analysis results reporting is a world where analysis results are machine-readable, easily navigable, and highly reusable. We envision the following:

- A logical model for describing analysis and results data
- Automated generation of machine-readable results data
- Improved navigation and reusability of analysis and results data
- Support for the storage, access, processing, and reproducibility of results data
- Traceability to the study protocol/SAP and to input ADaM data
- Open-source tools for designing, specifying, building, and generating analysis results data

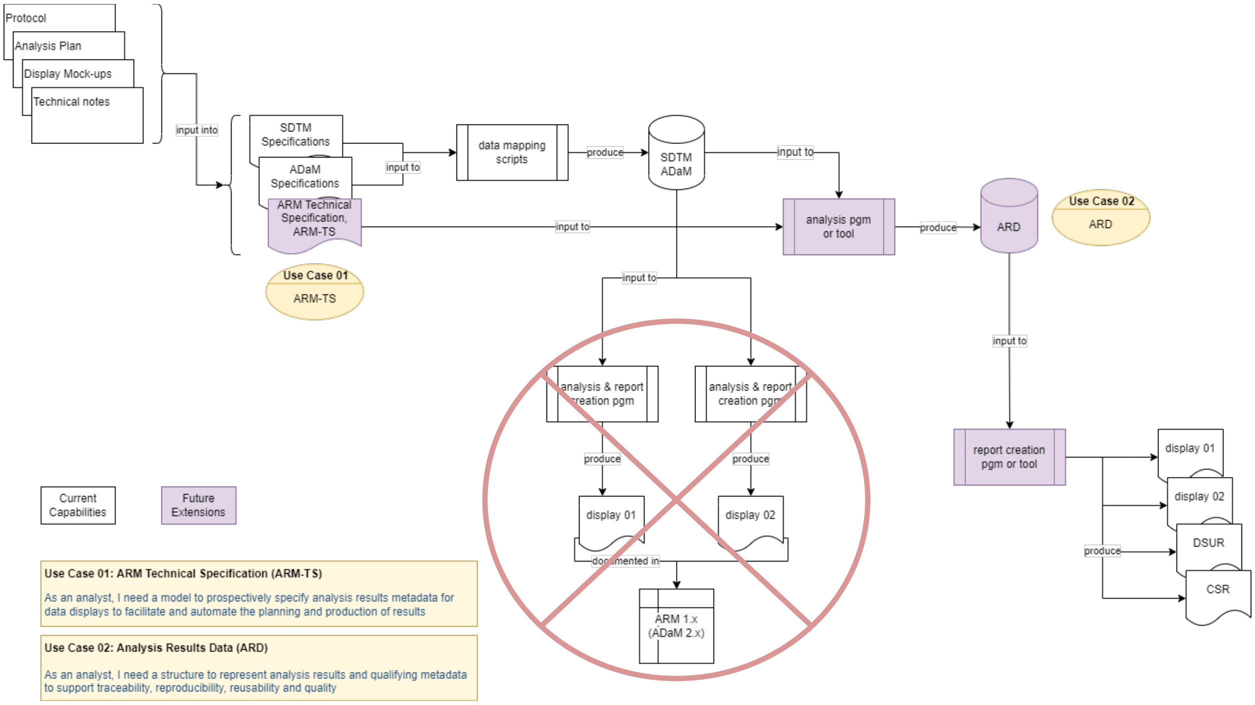
To achieve these goals, the ARS team has been working toward developing a logical model to fully describe analysis results metadata. This logical model will enable the implementation of an Analysis Results Metadata Technical Specification (ARM-TS) and an Analysis Results Data (ARD) framework. ARM-TS can be used to support automation, traceability, and the creation of data displays while the ARD framework will support reuse and reproducibility of results data.

New Approach

To address the current limitations of analysis results and associated metadata reporting, we are proposing a new workflow that shifts the focus from retrospective reporting to prospective planning. Specifically, we propose that end-users generate the Analysis Results Metadata Technical Specification (ARM-TS) prior to generating a display, rather than after the display has been created. This approach will allow for better planning and standardization of the analysis process, resulting in more consistent and traceable reporting.

The proposed workflow (Figure 2) involves several steps. First, the end-user will develop the ARM-TS, which will include metadata about the statistical methods, data sources, and displays to be generated (Figure 2, Use Case 1). Once the ARM-TS has been developed, the end-user will use it to generate an ARD, which will contain the results data needed to generate the display (Figure 2, Use Case 2). The ARD will be designed to support reuse and reproducibility of the results data, enabling more efficient and effective analysis reporting.

Finally, the machine-readable ARD serves as the ‘single source of truth’ capturing the analysis results metadata and results data in a standardized format. This ARD can then be used to generate displays for multiple reporting purposes, such as traditional analysis reporting for the clinical study report (CSR), in-text tables for the CSR, safety reporting, meta-analyses, dynamic applications, ClinicalTrials.gov, publications, and presentations. This streamlined approach ensures consistency and accuracy in the generation of displays across various deliverables, making it more efficient and reliable for reporting and communication of analysis results.



Use Case 01: ARM Technical Specification (ARM-TS)

As an analyst, I need a model to prospectively specify analysis results metadata for data displays to facilitate and automate the planning and production of results

Use Case 02: Analysis Results Data (ARD)

As an analyst, I need a structure to represent analysis results and qualifying metadata to support traceability, reproducibility, reusability and quality

Figure 2: Workflow with Future Extensions and Use Cases

Overall, this new workflow will enable end-users to generate analysis results metadata prospectively, with greater standardization, consistency, and traceability of analysis results reporting, enabling better decision-making and regulatory submissions. By shifting the focus from retrospective reporting to prospective planning, we believe that this approach will help to address many of the current limitations of analysis results reporting and support the development of more efficient and effective analysis standards.

In our vision for the future state, we anticipate the availability of open-source or community tools, such as the TFL Designer Community [1], in the industry. These tools will empower users to create machine-readable analysis metadata, which can automate the generation of analysis results data and displays. We also hope that such a tool can seamlessly integrate with existing analysis programs and report creation tools, enabling an end-to-end automation of the analysis and reporting process.

LINKML FOR LOGICAL MODEL:

Many of the same metadata components are needed both to create a prospective technical specification of analyses to be performed and to give context to reported results. The CDISC ARS team is therefore developing a single logical model comprising components needed to specify analyses, to represent contextualized results and to indicate how the results are displayed. The logical model is being developed using LinkML.

LinkML is an open-source schema development language and framework for generating machine-readable models [2]. The LinkML Generator framework generates downstream artifacts, including JSON-Schema, ShEx, RDF, OWL, GraphQL, and SQL DDL.

The logical model acts as a blueprint or a set of rules that describes how things should be organized and structured. In our case, it is being used to create a model that describes how analysis results data should be organized and structured. LinkML has allowed the creation of a standardized and consistent approach to describing analysis results data, enabling greater interoperability and collaboration across different stakeholders. Additionally, it has the added benefit of being able to convert them into various machine-readable formats such as JSON, YAML, OWL, or XML.

Furthermore, LinkML is a flexible and extensible tool, meaning that the model can be easily modified as needed to incorporate new data elements or requirements. By creating a machine-readable model, analysis results data can be more easily understood, shared, and reused by both humans and machines. LinkML will also support the development of validation rules that can be used to ensure the integrity and quality of the data, which is essential in our highly regulated pharmaceutical industry.

ANALYSIS RESULTS STANDARD MODEL:

An Analysis Results Standard model refers to the framework or structure used to store and organize the information related to an analysis result. This includes both the results data generated by the analysis and the metadata that describes the characteristics of that results data.

The metadata in this model includes information about the source of the results data, the method used for analysis, and any relevant parameters or queries that were used. This metadata is important for helping others understand and reproduce the analysis result, and it can also be used for quality control purposes.

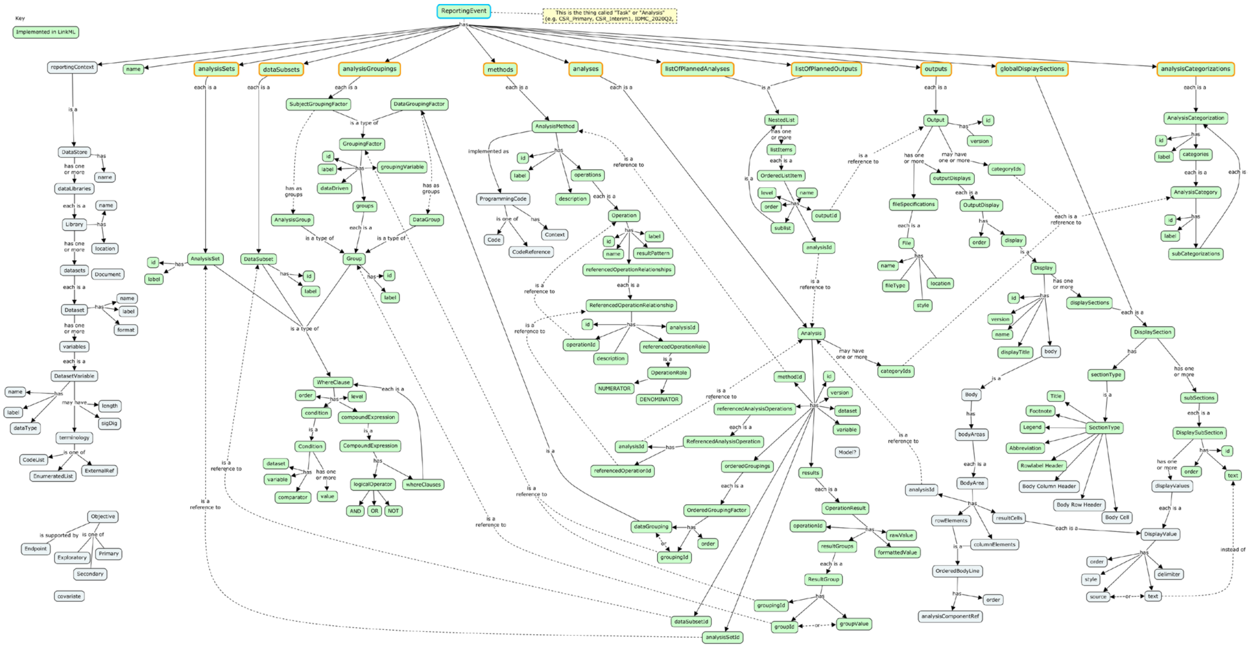


Figure 3: ARS Model [DRAFT] CMAP representation [3]

The model defines a set of classes and their attributes, which can be used to represent different aspects of the ARM-TS and ARD. The model includes classes such as "ReportingEvent", "Analysis", “AnalysisSet”, “AnalysisGroup”, "Output", and "GroupingFactor", among others. Each class has a set of attributes that define its properties, such as "id", "label", “condition”, “whereClauses”, "methods", "fileSpecifications", and so on.

The "ReportingEvent" class serves as the root of the model and includes many other classes and attributes that help to organize the data into different categories and groupings. Each reporting event represents a set of analyses and outputs created to meet a specific reporting requirement, such as a CSR or interim analysis. As shown in Figure 3, the ReportingEvent class has a set of attributes that can be used to define, itemize, and categorize the analyses and outputs needed for the reporting event:

- **analysisSets**:
  The analysis sets defined for the reporting event. Each analysis set, or subject population, is defined using the "AnalysisSet" class as a set of subjects whose data are to be included in the main analyses. This is as defined in the statistical section of the protocol.
- **dataSubsets**:
  Any defined data subsets that are used to restrict the records that are included in any given analysis (e.g., inclusion of only treatment-emergent adverse events in an AE summary analysis). Each data subset is defined using the "DataSubset" class.
- **analysisGroupings**:
  Characteristics used to subdivide the subject population (e.g., treatment, sex, age group). Each analysis grouping is defined using the "SubjectGroupingFactor" class and may either contain a set of specified groups (e.g., "Male" and "Female" for a sex grouping) or be defined as "data-driven", where the group values are obtained from the data (e.g., system organ class or preferred term for adverse events).
- **methods**:
  Defined methods used to analyze any analysis variable. Each method, which is defined using the AnalysisMethod class, may contain multiple individual operations. For example, a "summary of categorical variable" method might include separate operations for "count of subject by group" and "percent of subjects by group".
- **analyses**:
  The analyses defined for the reporting event. Each analysis is described, using the Analysis class, as a set of operations (i.e., a referenced "AnalysisMethod") performed on a specified analysis variable for any specified subset of data records (i.e., any referenced "DataSubset") for a given subject population (i.e., a referenced "AnalysisSet") which may be subdivided by one or more factors (i.e., any referenced "SubjectGroupingFactor" or "DataGroupingFactor"). The results of each analysis are contained within the "Analysis" class as a property of the analysis and each result is associated with references to all analysis components needed to give the result context (i.e., the operation used to calculate the result and a specific value for each grouping factors used for the analysis such as treatment, sex, age group, etc.)
- **outputs**:
  The outputs defined for the reporting event. Each output is defined using the Output class and may contain one or more defined displays.
- **listOfPlannedAnalyses**:
  A structured list of the analyses defined for the reporting event. The list is defined using the NestedList class, which allows referenced analyses (and outputs) to be organized into sections and sub-sections.
- **listOfPlannedOutputs**:
  An optional separate structured list of the outputs defined for the reporting event. The list is defined using the NestedList class, which allows referenced outputs (and analyses) to be organized into sections and sub-sections.
- **analysisCategorizations**:
  Any categorizations used to tag analyses or outputs. For example, a "Type of analysis" categorization could be created to tag defined analyses or outputs as "Subject Population", "Safety" or "Efficacy", and/or an "Estimand Analysis Type" categorization could be created to tag analyses as "Primary Estimator", "Sensitivity Analysis", "Supplementary Analysis". Each categorization is defined using the "AnalysisCategorization" class.

Due to the complex nature of the model, we have provided a video recording as well as the copy of the presentation that can be accessed in the public domain following the links below. All the files referenced in the presentation can be accessed from the CDISC Analysis Results Standard public repository on GitHub [3]. If the hyperlinks are not functioning, please reach out to the authors using the contact information provided.

Video Recording	Slide Deck	GitHub Repo

While the logical model is necessarily complex, due to the complexity of analysis, we expect that the existence of the model will support the development of open-source tools to facilitate the creation of prospective analysis specification metadata, to re-represent the metadata in any necessary human- or machine-readable format, and ultimately to automate the process for generation and display of analysis results.

CONCLUSION

The ARS team has been actively working on creating new analysis results standards. We have discussed the current state, future state, and logical model for these standards in this paper. We are excited to announce that the first version of the standards will be released in the summer of 2023, which will include a logical model to support ARM-TS/ARD and four common safety examples based on team-developed tables for demographics, adverse events, and vital signs.

Moving forward, we plan to expand the use cases of the standards, develop APIs for the extraction of examples from the CDISC Library, establish conformance rules, and standardize terminology. Additionally, we will be conducting public hackathons to encourage the industry to build open-source tools around these standards and promote adoption.

We believe that these efforts will not only enhance the interoperability and reusability of analysis results data but also promote innovation in the pharmaceutical industry. We look forward to collaborating with industry and regulatory bodies to make these standards a reality.

REFERENCES

1. (TFL Designer GitHub, 2023): https://github.com/bhavinbusa/tfldesigner
2. (The Linked Data Modeling Language: A framework for describing and integrating rich biomedical data, 2022): https://www.slideshare.net/cmungall/linkml-intro-july-2022pptx
3. (CDISC Analysis Results Standard GitHub, 2023): https://github.com/cdisc-org/analysis-results-standard

Additional Resources:

4. Busa, Bhavin; LeRoy, Bess; 'Pre-launching CDISC Analysis Results Standards' at PHUSE US Connect (Mar 2023) Presentation Link, Video Link
5. Busa, Bhavin; LeRoy, Bess; Marshall, Richard; 'Hands-on Workshop: Deep-dive Review and Testing of CDISC Analysis Results Standards Logical Model and Schema' at PHUSE US Connect (Mar 2023). Workshop files available on ARS GitHub.
6. Wallendszus, Karl; Frenzel, Hansjoerg; 'CDISC Analysis Results Standard – Update and Progress' at PHUSE EU Connect (Nov 2022). Paper Link, Presentation Link
7. Busa, Bhavin; LeRoy, Bess; Miskell Andrew; 'CDISC Analysis Results Standards – Approach and Development' at CDISC US Interchange (Oct 2022)

8. Barros, J.M., Widmer, L.A., Baillie, M. et al. Rethinking clinical study data: why we should respect analysis results as data. Sci Data 9, 686 (2022). https://doi.org/10.1038/s41597-022-01789-2

9. PHUSE White Paper “General Output Tips and Considerations”, Doc ID: WP-034, Version 1.0, Aug 2020

10. CDISC Analysis Results Standard landing page: https://www.cdisc.org/standards/foundational/analysis-results-standards

ACKNOWLEDGMENTS

We would like to extend our heartfelt thanks to the entire CDISC Analysis Results Standard team, whose hard work, dedication, and expertise have been instrumental in developing this new analysis results standard. Your contributions have been invaluable in helping to improve the standards and underlying model, and we are grateful for your collaboration and commitment to advancing the field.

CONTACT INFORMATION

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

Bhavin Busa
ARS Product Owner & Co-Lead, CDISC
Principal & Co-founder, Clymb Clinical
bhavin@clymbclinical.com

Richard Marshall
Principal Data Modeler
rmarshall@accuratesystems.co.uk

Bess LeRoy
ARS Co-Lead
Head of Standards Innovation, CDISC
bleroy@cdisc.org