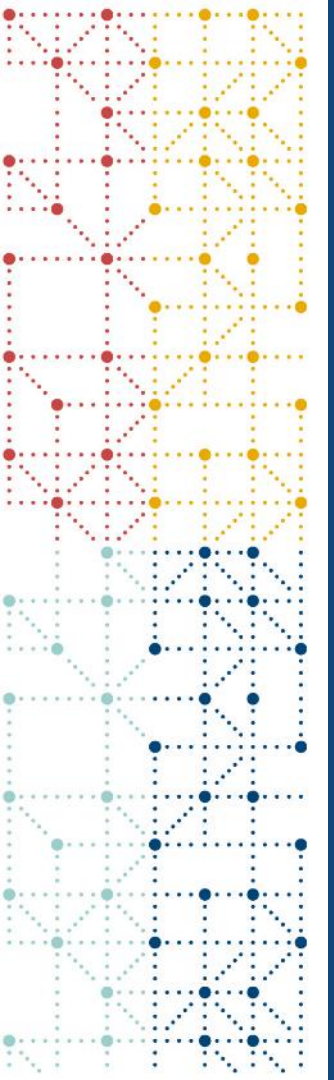
cdisc

2023
JAPAN
INTERCHANGE
TOKYO | 10-11 JULY

CDISC/PHUSE Collaboration

Presented by ***Yoshiko Kitagawa***
Associate Director,
CDISC Standard, Data Science
ONO Pharmaceutical Co., Ltd

Supported by ***Chris Price, Working Groups Director, PHUSE***



Yoshiko Kitagawa

Associate Director,

CDISC Standard, Data Science

ONO Pharmaceutical Co., Ltd

20 years' experience of statistical programmer/Standards programmer within the pharmaceutical industry - specializing in CDISC Submission data with cross-cultural experiences thought tasks working for the companies in Japan and the UK.

PHUSE WG

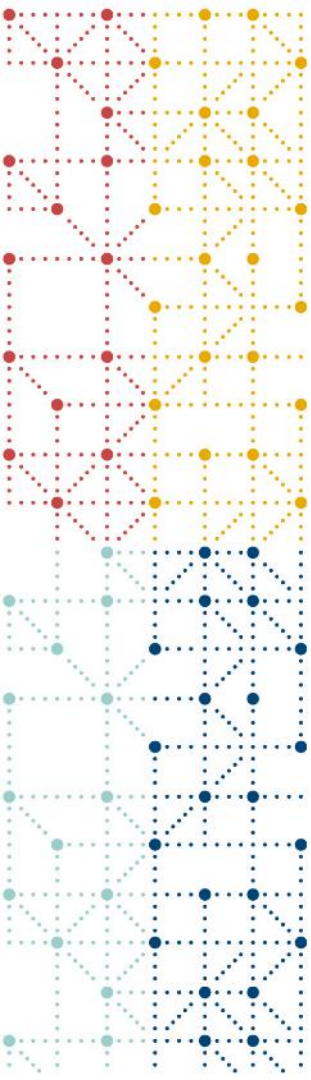
- PHUSE.global WG >>
- - 2021: FDA and PMDA Study Data Submission Distinctions
- 2017 - Current:
- SDTM ADaM Implementation FAQ,
  Clinical Integrated Study Data & Analysis Data Reviewer's Guide,
  …
- PHUSE.JP >>
- 2014: PHUSE Japan SDE WG
- 2021: PHUSE Japan SDE Chair
- -Current: Electronic Data Submission in Japan

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not reflect the official policy or position of PHUSE, CDISC and ONO Pharmaceutical.*

Agenda

- CIDSC and PHUSE
- Dataset-JSON Pilot Project
- Ongoing PHUSE/CDISC projects
  - SDTM/ADaM Implementation FAQ
  - Implementation of Estimands using data standards
  - SEND Industry Feedback Survey
  - Harmonization of SEND Implementation to Enable Historical Control Data Analysis
  - SEND Implementation User Group
- PHUSE/CDISC Events



CDISC and PHUSE

About CDISC

- **Clinical Data Interchange Standards Consortium**
- https://www.cdisc.org/about

About PHUSE

The Global Healthcare Data Science Community

Our Mission

Sharing ideas, tools and standards around data, statistical and reporting technologies to advance the future of life sciences.

Who We Are and What We Do

PHUSE is an independent, not-for-profit organisation run by a worldwide team of volunteers. We are a global community and platform for the discussion of topics encompassing the work of data managers, biostatisticians, statistical programmers, data scientists and eClinical IT professionals. PHUSE has become the industry voice to regulatory agencies and standards organisations such as the FDA, EMA & CDISC.

From:

https://phuse.global/About_PHUSE

- **PHUSE/CDISC Collaboration Project**

- Dataset-JSON Pilot Project

Limitations of SAS V5 XPORT Format

Data File Format	Storage	Content	Extensibility
Limited variable type Limited to US ASCII encoding 8-character variable names 40-character labels 200-character field widths	Inefficient use of storage space Inability to compress datasets leads to file logistical issues (e.g., splitting datasets)	Lacks a robust metadata layer Only works for 2-dimensional data structures	Not extensible

From the PHUSE *Transport for the Next Generation* (2017) White Paper

What is Dataset-JSON and Advantages

What is JSON?
An open standard file format and data interchange format that uses human-readable text to store and transmit data objects consisting of attribute–value pairs and arrays

What is Dataset-JSON?
A dataset exchange standard for exchanging tabular data leveraging JSON designed to meet the regulatory submission needs and eliminating limitations of legacy formats

Dataset-JSON is…

- Part of the ODM 2.0 standard
- An open-source MIT license
- Schema supporting any tabular format
- Extensible to support integrated metadata and new use cases
- Linked to Define-XML for complete metadata
- Integrated with CORE for conformance checking

Dataset-JSON advantages…

- Based on the JSON standard used worldwide
- Open-source and truly human readable
- Same or smaller file sizes relative to current required format
- Remove variable naming, width, or format limitations
- Simple transformation to/from SAS data

COSA Dataset-JSON Hackathon 2022

- Dataset-JSON Hackathon open-source solutions available in the COSA Repository Directory
- Solutions created include:
  - Conversion to and from different dataset formats
  - Dataset browsers / viewers
  - Methods for handling large datasets
  - RESTful Web Services
- Overall Impressions of Dataset-JSON:
  - Works as a general data exchange
  - Works as a general dataset format
  - Works with web-based APIs
  - Works with a wide-range of programming languages and technology stacks
  - Simple to process
  - Easy to transform into SAS datasets, R or Python dataframes, and CSV
  - File sizes smaller than SAS XPORT v5 and Dataset-XML
  - A language, platform independent data exchange format

Language	# Solutions
R	5
SAS	4
Python	5
JavaScript	4
Java	1
Swift	1
XSLT	1

https://cosa.cdisc.org/

Dataset-JSON Pilot

Milestone 1: Short Term

- Pilot submissions using JSON format with existing XPT ingress/egress to carry the same data
- Same content, different suitcase, no disruption to business process on either side
- In parallel, evaluate with FDA how their toolset can support JSON format and identify tool upgrade roadmap

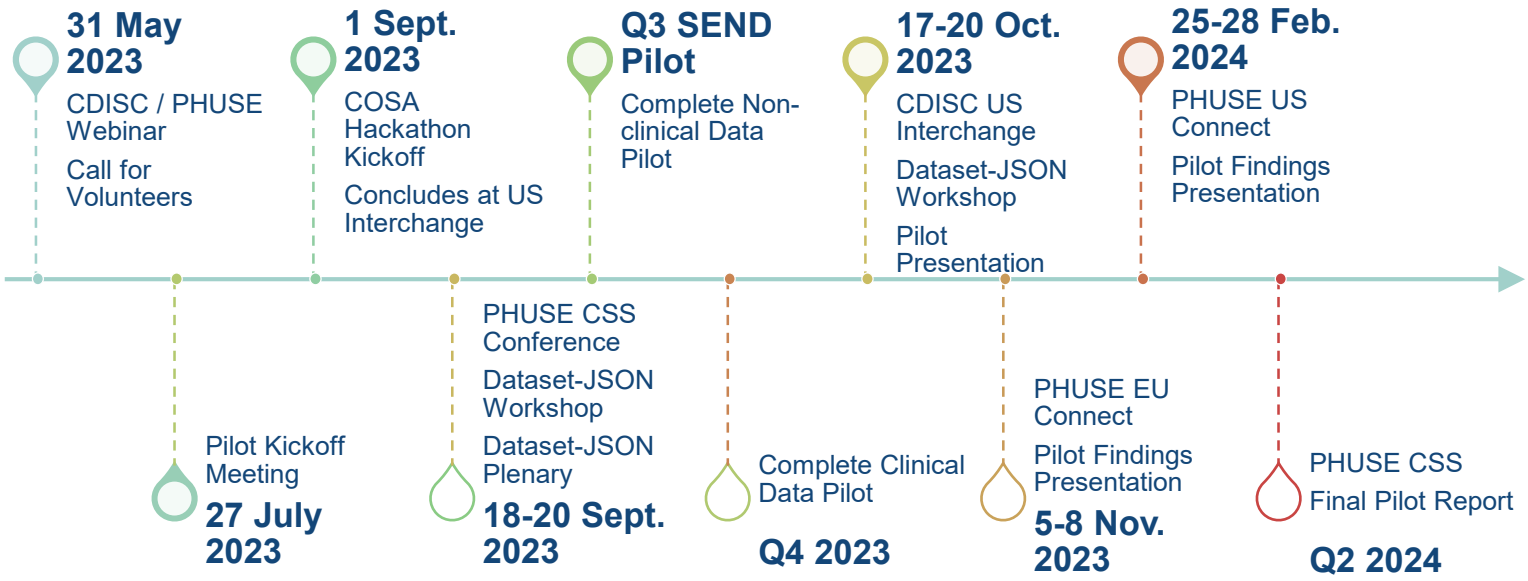
➔ **Success Criteria: Accept Dataset-JSON as a transport format option (in addition to existing XPT format)**

Milestone 2: Development of future strategy

- Evaluate how current and future industry standards can benefit without XPT limitations
  - e.g. Variable names > 8, labels > 40, data > 200
- Evaluate combining metadata with data
  - e.g. Define-XML / Define-JSON based
- Enhanced conformance rules
- Collaborate with FDA to develop plan to retool their environment to natively consume JSON

➔ **Success Criteria: accept advanced Dataset-JSON as the only transport format option and deprecate XPT**

Dataset-JSON Pilot: draft Timeline



Non-clinical & Clinical Data Pilot Overview

Non-clinical Data Pilot

- Based on recently concluded non-clinical, SEND, pilot
- Convert datasets used in the pilot
- Collaborate with the PHUSE Nonclinical Topics WG
- Planning to conclude the pilot in Q3 of 2023

Clinical Data Pilot

- Ask sponsors to test conversions on in-house data
  - No external data transfers involved
  - Conversion software will be provided
- Sponsors will review the conversion process and provide feedback
- CDISC Pilot datasets will be sent to FDA to test the gateway
- Planning to conclude the pilot by EOY 2023

➔PHUSE – CDISC project will consolidate pilot feedback
➔FDA is expected to participate, discussions are ongoing

COSA Dataset-JSON Hackathon II 2023

- Primary objective: Create a draft API specification for Dataset-JSON
- Secondary objective: Proof-of-concept implementations to demonstrate and test the API specification
- Virtual hackathon
  - May have a workshop to wrap-up the API specification at the US Interchange
  - Will read out the results of the hackathon during the Interchange
- Dates: September 1 – October 19

PHUSE CSS and Dataset-JSON

18 September

PHUSE/CDISC Workshop

- Gain an understanding of Dataset-JSON and prepared to participate in the pilot

18 September

Plenary Session

- Collaboration driving Technology Adoption (including Dataset-JSON and Open Source)

19 – 20 September

Working Sessions

- Collaborative sessions bring together Emerging Trends & Technology, Optimizing the Use of Data Standards and Nonclinical Topics Working Groups

Other Dataset-JSON Conference Activities

2023 CDISC US Interchange	• 2-hour hands-on workshop prior to the conference • Dataset-JSON Pilot presentation
2023 PHUSE EU Connect	• Dedicated workshop • Focused on engaging community
2024 PHUSE US Connect	• Dedicated workshop • Presentation to cover pilot findings and next steps
2024 PHUSE CSS	• Presentation to cover final pilot report

- Ongoing PHUSE/CDISC projects
  - SDTM/ADaM FAQ
  - Implementation of Estimands using data standards
  - SEND Industry Feedback Survey
  - Harmonization of SEND Implementation to Enable Historical Control Data Analysis
  - SEND Implementation User Group

Ongoing PHUSE/CDISC Projects
Optimizing the Use of Data Standards

SDTM ADaM Implementation FAQ

- Provide a forum and subsequent knowledge-base (FAQ) to address common challenges amongst SDTM and ADaM implementers and consumers as well as those related to data submissions requirements by various regulatory agencies around the world

Best Practices in Data Standards Implementation Governance

- Best practices on how sponsors/CROs can govern the implementation of data standards
- How to build and maintain content in internal metadata standards repositories
- How to consistently extend data standards for individuals projects/studies (and how to maintain these extensions)
- What should be governed centrally vs permitted on studies?
- How to feedback/communicate to/with SDOs

Ongoing PHUSE/CDISC Projects
SDTM ADaM Implementation FAQ

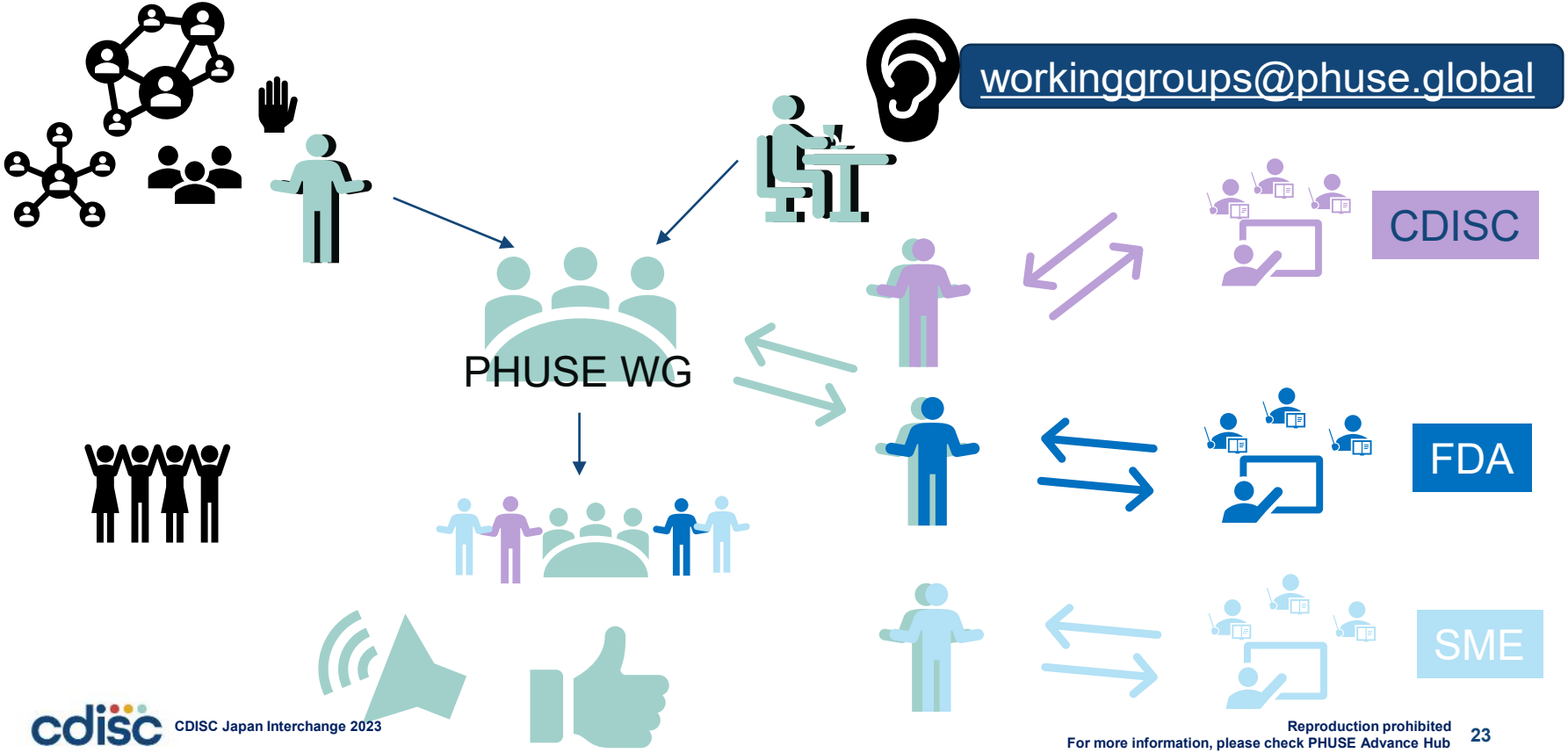
PHUSE/CDISC Project

- The Standards Implementation Nuances sessions at the March North America 2016 CSS and June EU 2016 CSS surfaced various common challenges amongst SDTM and ADaM implementers and consumers. In addition, there were many questions discussed related to data submissions requirements by various regulatory agencies around the world (e.g. FDA, PMDA, China FDA). It became very clear that the Industry is in need of a forum and subsequent knowledge-base (FAQ) to address these challenges. This project team was formed in June 2016 to collaborate with Subject Matter Experts (SMEs) from the industry, CDISC, and the FDA.

Goals

- Collect frequently asked questions (FAQs) from Industry. If you have a question for the team, email workinggroups@phuse.global.
- Assess the appropriateness of a question, develop & review a response, collaborate with CDISC/FDA for clarity if required.
- Publish the FAQ and responses on Advance Hub database for helpful implementation/strategy information.

Ongoing PHUSE/CDISC Projects
SDTM ADaM Implementation FAQ



Ongoing PHUSE/CDISC Projects
SDTM ADaM Implementation FAQ

Question	Teams Collective Response
What kind of information about a subject with multiple screenings needs to be submitted to the FDA?	**PHUSE Team Response: 11 March 2022** While the FDA's Study Data Technical Conformance Guide (v4.8.1, October
These questions are primarily going out to the sub-team that worked on the Best Practices for Submission of Event Adjudication Data White Paper. The White Paper provided very useful tips on how to map adjudicated data to the new custom SDTM domain of EA. The following are the follow-up questions to this White Paper.	**PHUSE Team Response: 14 July 2020** CDISC SDS informed that the adjudication project is under consideration and may be in the future SDTMIG (beyond SDTMIG v3.4). Submitting EA as a custom domain is allowed by the current SDTMIG. The proposed domain in the White Paper is based on previous
How should the sex of transgender patients be collected and analysed in clinical trials? Should the sex at birth be collected only or should the gender preference also be collected? Which laboratory normal ranges should be assigned to transgender patients' laboratory test results? How does hormone therapy affect data collection and/or analysis for transgender patients?	**PHUSE Team Response: 30 June 2022** The CDISC CDASH team is currently working to publish either an updated guidance or white paper **planned for** 2025 on recommendations on capturing the sex for transgender patients. In the draft version, the recommendation would be to collect a two-stage question (note that the controlled terminology and collection text are a draft stage and not finalised): 1. "Sex at Birth" (Male | Female | Don't know | Prefer not to answer) and 2. "Sexual Identity" (Male | Female | Intersex | Transgender | ⋯ | Don't know | Prefer not to answer | Self-describe). In the interim, each sponsor should determine how the data should be collected. It is recommended to provide clarity on the definition of each question, perhaps within the CRF Completion Guidelines. For example, does Sex at Birth pertain to sex stated on the birth certificate, and how to complete the data entry if a patient does not have a birth certificate. The following articles may be reviewed to determine how hormone therapy affects laboratory results and, in general, analysis for transgender subjects: 1. "Common Hormone Therapies Used to Care for Transgender Patients Influence Laboratory Results", Humble, R. et al, 2018, *American Association for Clinical Chemistry*. 2. "Interpreting Laboratory Results in Transgender Patients on Hormone Therapy", Roberts, T. et al, 2014, *The American Journal of Medicine*. 3. "Impact of Hormone Therapy on Laboratory Values in Transgender Patients", SoRelle, J. et al, 2019, *Clinical Chemistry*. 4. "Approach to Interpreting Common Laboratory Pathology Tests in Transgender Individuals", Cheung, A. et al, 2021, *The Journal of Clinical Endocrinology & Metabolism*. 5. "Endocrine Treatment of Gender-Dysphoric/Gender-Incongruent Persons: An Endocrine Society* Clinical Practice Guideline", Hembree, W. et al, 2017, *The Journal of Clinical Endocrinology & Metabolism*.

Ongoing PHUSE/CDISC Projects
SDTM ADaM Implementation FAQ

Question:

What kind of information about a subject with multiple screenings needs to be submitted to the FDA?
The FDA Study Data Technical Conformance Guide mentioned…

Response:

…

The Multiple Subject Instances Team at CDISC is currently working on creating the new DC domain that will address this issue. For each USUBJID, this domain will contain multiple entries for each time that the subject screened or enrolled into the study. The SUBJID value will reflect the subject identifier for that time of participation. The Multiple Subject Instances Team at CDISC is also recommending the SUBJID to be included in the parent domain as a permissible variable, …

Ongoing PHUSE/CDISC Projects
Implementation of Estimands Using Data Standards

ICH E9(R1)

- Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials
- Finalized in November 2019 (Step 4)
- Has been or is in the process of being adopted by Health Authorities
- Covers the important multidisciplinary considerations relating to the implementation of the ICH E9(R1) estimands framework for clinical trial planning, design, conduct, analysis and interpretation
- The technical implementation in the data flow was not in scope
- Additional guidance for the implementation of the estimands framework in the data flow is necessary

PHUSE/CDISC Project

- Developing a White Paper to provide recommendations and best practices to implement the estimands framework in data standards
- Consultation and feedback loop to CDISC teams responsible for standards development on proposed implementation

Ongoing PHUSE/CDISC Projects
Implementation of Estimands Using Data Standards

Data Collection	• Accuracy and Granularity • Sponsors should assess study designs
Codelists	• Proposals submitted to for new terminology • Recommendations for new terms
SDTM	• Estimands framework has no impact • Follow SDTM IG & Conformance Rules
cSDRG	• Section for Intercurrent Events • Define, collection and mapping

Ongoing PHUSE/CDISC Projects
Implementation of Estimands Using Data Standards

ADICE

- Consistent documentation of all intercurrent events
- Support harmonized workflows

New ADaM Variables

- ADSL: New estimand analysis set flag
- BDS: New record level data point flag and intercurrent event traceability variables

Guidance

- Building upon existing ADaM-IG that already addresses analysis features (estimations).

Ongoing PHUSE/CDISC Projects
Nonclinical Topics Working Group

SEND Implementation User Group	• Resources to help navigate SEND implementation • An opportunity to ask questions, receive answers and to discuss and learn more • Investigate standard implementation challenges encountered during use by industry • Provide feedback and recommendations to CDISC for consideration in standards training and development
SEND Industry Feedback Survey	• Execute an annual survey to collect data that enables detection of issues and trends • Provide SDOs with actionable feedback regarding improvements to their standards • Enable all with extensive data management experience, an opportunity to provide actionable recommendations • Help Sponsors better leverage their SEND investments
Harmonization of SEND Implementation to Enable Historical Control Data Analysis	• Recommend approaches to better enable analysis of historical control data. • Includes creation of new variables, controlled terminology, preferred terms, reference lists and/or analysis strategies to enable cross study analysis. • Developing a solution framework for variability to enable more efficient routine analysis of warehoused SEND data

- **PHUSE/CDISC events & collaboration**

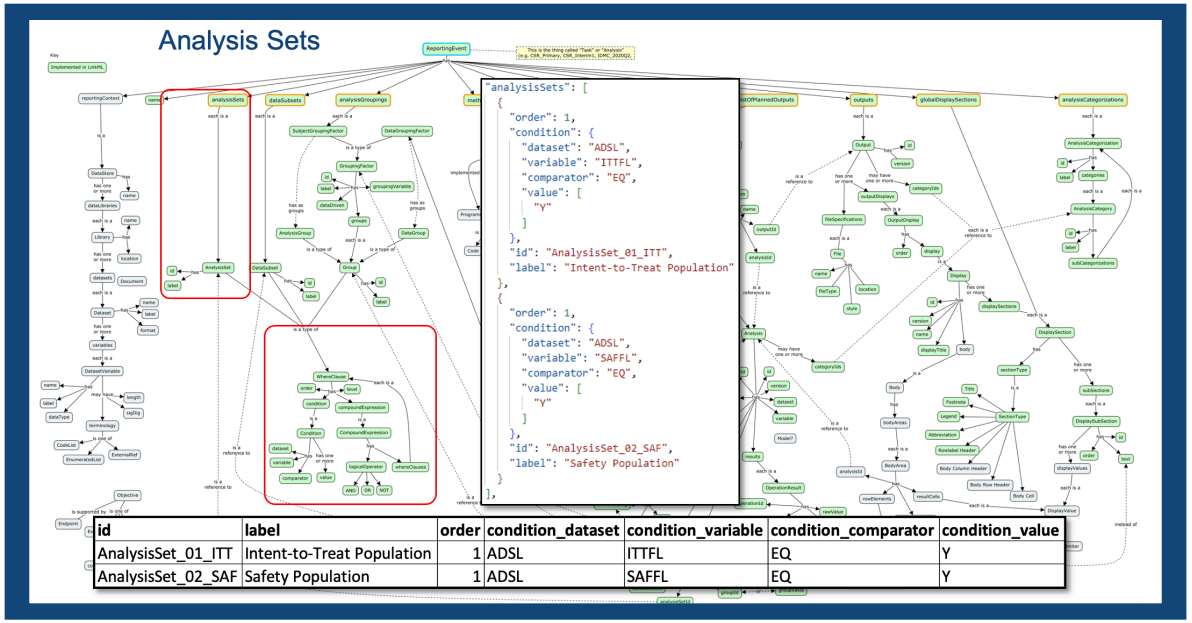
PHUSE & CDISC Collaboration

Events	• Mutual representation on events • Engage with Community directly
Communication	• Joint communication initiatives
Community	• Active participation in volunteer groups • Projects and Community often overlap
Workshops	• Use of CDISC Workshops at PHUSE events • Drive adoption – engage PHUSE Community

PHUSE US Connect 2023: Analysis Results Standard Workshop

- CDISC's Analysis Results Standard aims to leverage analysis results metadata to help drive automation and to support storage, access, processing and reproducibility of results

- Over 70 attendees participated in a hands-on workshop that took a deep dive into the logical model and safety examples



id	label	order	condition_dataset	condition_variable	condition_comparator	condition_value
AnalysisSet_01_ITT	Intent-to-Treat Population	1	ADSL	ITTFL	EQ	Y
AnalysisSet_02_SAF	Safety Population	1	ADSL	SAFFL	EQ	Y

PHUSE JP

- Working Groups
  - Data Transparency
  - Data Visualisation & Open Source Technology
  - Emerging Trends & Technologies
  - Nonclinical Topics
  - Optimizing the Use of Data Standards
    - Management of ODS Regulatory Referenced Deliverables
    - Bioresearch Monitoring (BIMO) Frequently Asked Questions Forum
    - (BIMO) Bio-research Monitoring Data Reviewers Guide
    - Implementation of Estimands (ICH E9 (R1)) using Data Standards
    - Best Practices in Data Standards Implementation Governance
    - Clinical Integrated Study Data & Analysis Data Reviewer's Guide
    - SDTM ADaM Implementation FAQ
    - **Electronic Data Submission in Japan**

Thank You!