

Useful Tools from the Computational Sciences Symposium

Sandra Minjoe, Accenture Life Sciences
John Brega, PharmaStat



PharmaSUG Single Day Event
San Francisco Bay Area

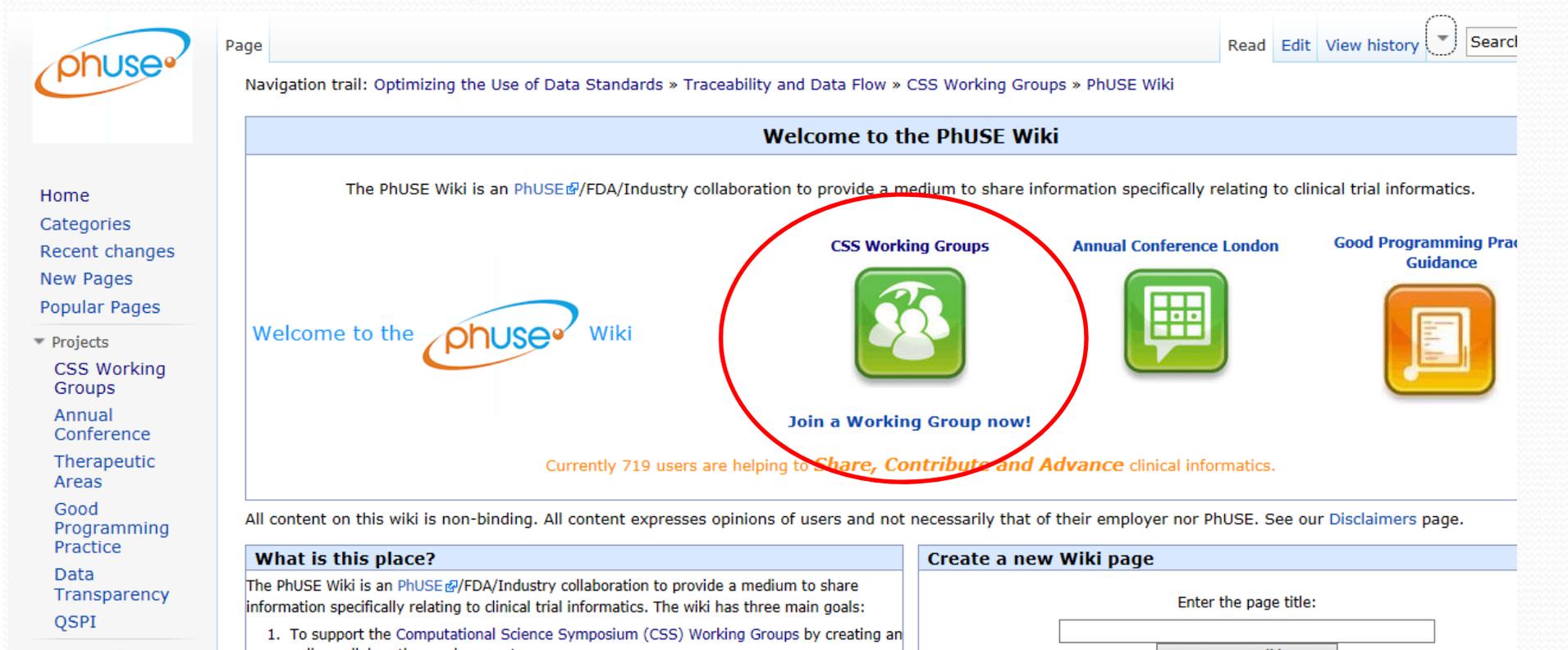
2015-02-10

What is the Computational Sciences Symposium?

- CSS originally formed to help FDA CDER Computational Sciences Center with technology
- Addresses topics not well covered by CDISC
- Annual face-to-face working meeting in March but most work is done via telecons
- PhUSE maintains wiki where CSS deliverables are available to all

Computational Sciences Symposium Wiki Navigation

- phusewiki.org



The screenshot shows the PhUSE Wiki homepage. At the top left is the PhUSE logo. Below it is a navigation menu with links for Home, Categories, Recent changes, New Pages, Popular Pages, and Projects. The Projects section is expanded, showing links for CSS Working Groups, Annual Conference, Therapeutic Areas, Good Programming Practice, Data Transparency, and QSPI. The main content area features a 'Welcome to the PhUSE Wiki' header, followed by a paragraph describing the wiki as a PhUSE/FDA/Industry collaboration. Below this are three featured sections: 'CSS Working Groups' (highlighted with a red circle), 'Annual Conference London', and 'Good Programming Practice Guidance'. Each section has an icon and a call to action. The 'CSS Working Groups' section includes the text 'Join a Working Group now!' and 'Currently 719 users are helping to Share, Contribute and Advance clinical informatics.' At the bottom, there are two boxes: 'What is this place?' and 'Create a new Wiki page'.

Page Read Edit View history Search

Navigation trail: [Optimizing the Use of Data Standards](#) » [Traceability and Data Flow](#) » [CSS Working Groups](#) » [PhUSE Wiki](#)

Welcome to the PhUSE Wiki

The PhUSE Wiki is an [PhUSE](#)/[FDA](#)/Industry collaboration to provide a medium to share information specifically relating to clinical trial informatics.

Welcome to the  Wiki

- CSS Working Groups**

Join a Working Group now!
Currently 719 users are helping to **Share, Contribute and Advance** clinical informatics.
- Annual Conference London**

- Good Programming Practice Guidance**


All content on this wiki is non-binding. All content expresses opinions of users and not necessarily that of their employer nor PHUSE. See our [Disclaimers](#) page.

What is this place?

The PhUSE Wiki is an [PhUSE](#)/[FDA](#)/Industry collaboration to provide a medium to share information specifically relating to clinical trial informatics. The wiki has three main goals:

1. To support the Computational Science Symposium (CSS) Working Groups by creating an online collaboration environment

Create a new Wiki page

Enter the page title:

CSS Wiki Navigation

Navigation trail: [Optimizing the Use of Data Standards](#) » [Traceability and Data Flow](#) » [PhUSE Wiki](#) » [CSS Working Groups](#)

CSS Working Groups

(Redirected from [FDA Working Groups](#))

Instructions

[\[edit\]](#)

This wiki is available to document the progress of the PhUSE Computational Science (CS) working groups and projects. The wiki enables quick, easy and transparent online collaboration. Working groups each have a main page which is the starting page for all Working Group activity.

If you are new to wikis you might want to **check out the help section**. To get started click on your working group below.

For instructions on how to subscribe to the working group/project email list or contact a working group leadership team [click here](#)

The Working Groups are governed by the [Computational Science Steering Committee](#) which provides oversight to the Working Groups.

PhUSE CS Working Groups

[\[edit\]](#)

[CS Dashboard on www.phuse.eu](#)

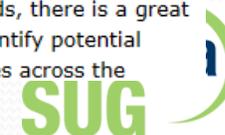
[CS Catalog of Deliverables on www.phuse.eu](#)

[CS Catalog of Deliverables Under Review on www.phuse.eu](#)

Optimizing the Use of Data Standards The development and adoption of data standards over the last decade has shown significant promise in improving efficiencies in the product submission and review process. However, there have also been gaps, issues and challenges in the interpretation and use of the standards. This group will identify specific gaps preventing FDA and industry from optimizing the use of standards and collaborate to close those gaps.

- **NEW:** A new Project within the Optimizing the Use of Data Standards is getting started to develop a templates for a **Study Data Standardization Plan..** Visit the Wiki page for more information on how you can participate on this new project.
- **ADRG and SDRG Work Packages Referenced in Final FDA Study Data Technical Conformance Guide, Dec 2014**

Development of Standard Scripts for Analysis and Programming With the development and implementation of industry data standards, there is a great opportunity to develop standard reporting across industry and to support the needs of FDA medical and statistical reviewers. This working group will identify potential standard scripts for data transformations and analyses across an within therapeutic areas. The goal will be to begin the process of standardizing analyses across the



CSS Wiki Navigation

Optimizing the Use of Data Standards

Working Group Overview

[\[edit\]](#)

The development and adoption of data standards over the last decade has shown significant promise in improving efficient delivery of data to support drug product and device submissions as well as the review process. However, there have also been gaps, issues and challenges in the interpretation and use of data standards. This working group will identify specific gaps that prevent FDA and industry from optimizing the use of data standards. This working group will collaborate to close those gaps. On this page you will find information on Current and Completed projects.

Contents [\[hide\]](#)

- 1 Working Group Overview
- 2 Leadership Team
- 3 Current Projects
- 4 Completed Projects
 - 4.1 Reviewer's Guide Work Packages referenced in FDA Study Data Technical Conformance Guide v2.0, Dec2014
- 5 Projects on Hold
- 6 Working Group Meeting Materials
- 7 Legacy Projects/Information

Leadership Team

[\[edit\]](#)

Name	Role	Organization	E-mail
Dhananjay Chhatre	Industry Lead	Booz Allen Hamilton	Chhatre_Dhananjay (at) bah.com
Michael Brennan	Steering Committee Liaison	Janssen R&D	MBrenna3 (at) its.jnj.com
Jeff Conant	Group Coordinator	Vertex	Jeffrey_Conant (at) vrtx.com
Jingyee Kou	FDA Liaison	FDA	jingyee.kou (at) fda.hhs.gov
Steve Wilson	FDA Liaison	FDA	Stephen.Wilson (at) fda.hhs.gov

Current Projects

[\[edit\]](#)

Traceability and Data Flow - Traceability challenges arise during the data flow process, for example when raw study level data is converted to SDTM after the fact, while analysis data sets and the study report trace back to the original raw data source. This project discusses and defines traceability considerations and best practices for both study level and integrated data set conversion, and for a variety of different data flow scenarios. Several white papers and a summary of traceability references have already been developed.

Best Practices For Data Standards Implementation - Development of a recommended set of best practices for optimizing implementation of data standards is the focus

CSS Wiki Navigation

Traceability and Data Flow

Project Overview

[edit]

It is always a challenge for the results in the clinical study report to be able to trace back to the original raw data source. The traceability challenges intensify when raw data is converted to SDTM after the fact, while analysis datasets and the study report trace back to the original raw data source. This project will discuss and define traceability considerations and best practices for study level dataset and integrated datasets conversion for a variety of different data flow scenarios.

Project Deliverables

[edit]

The following deliverables are available for review:

If you have any questions or comments, please contact the deliverable lead - listed in the Objectives and Timelines table

[Summary of Traceability References](#)

[Traceability: Current State Analysis](#)

[Preliminary Recommendations for Traceability Documentation using Define-XML 2.0](#)

[Traceability: Best Practices for Basic Linear Data Flow **NEW VERSION: Now includes Recommendations**](#)

[White Paper: Study Level Traceability in a Non Linear Data Flow **NEW: Updated October 2014**](#)

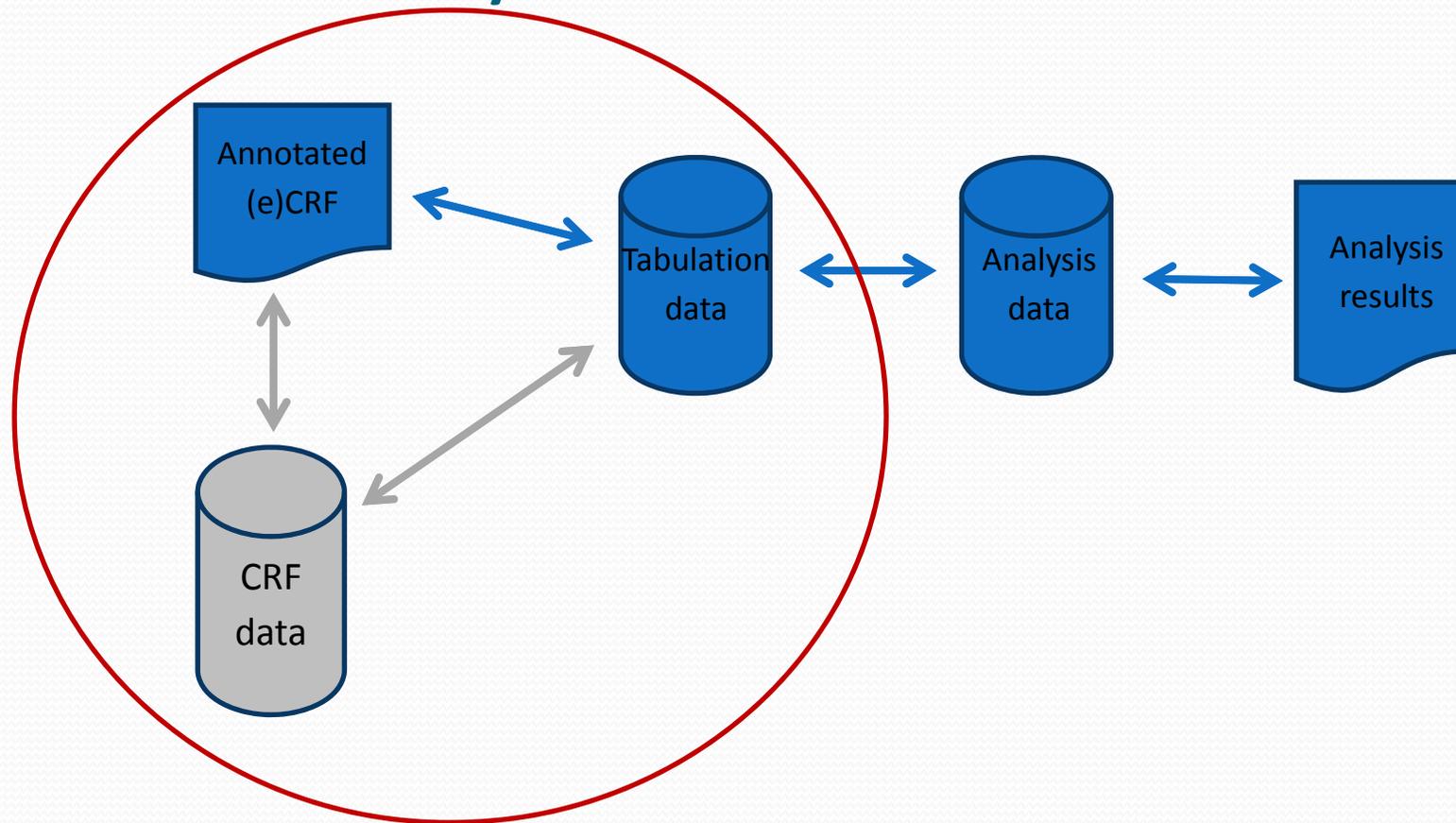
Traceability and Data Flow Project

- Traceability goal:
 - Follow from the results in the clinical study report back to the original raw data source
- Project attempts to discuss/define traceability considerations and best practices for
 - Study level dataset
 - Integrated datasets
- Can be used with or without CDISC structures

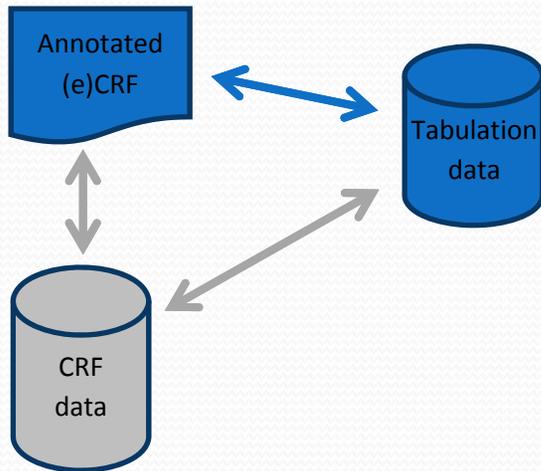
Traceability & Data Flow Projects Completed

- Summary of Traceability References
- White Papers
 - Traceability: Current State Analysis
 - Preliminary Recommendations for Traceability Documentation using Define 2.0
 - Traceability: Best Practices for Basic Linear Data Flow
 - Study Level Traceability in a Non Linear Data Flow

Traceability in Clinical Trials Data Flow

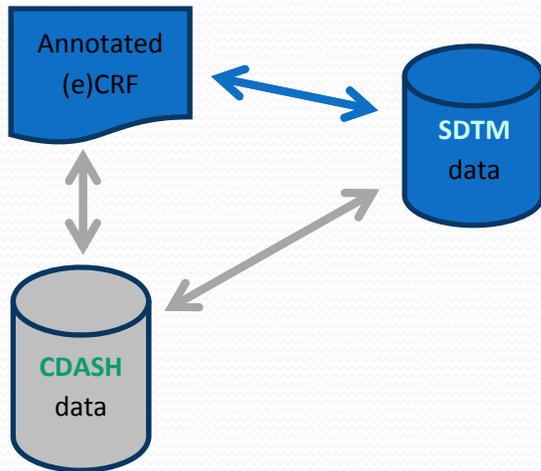


Traceability between Data Collection and Tabulation



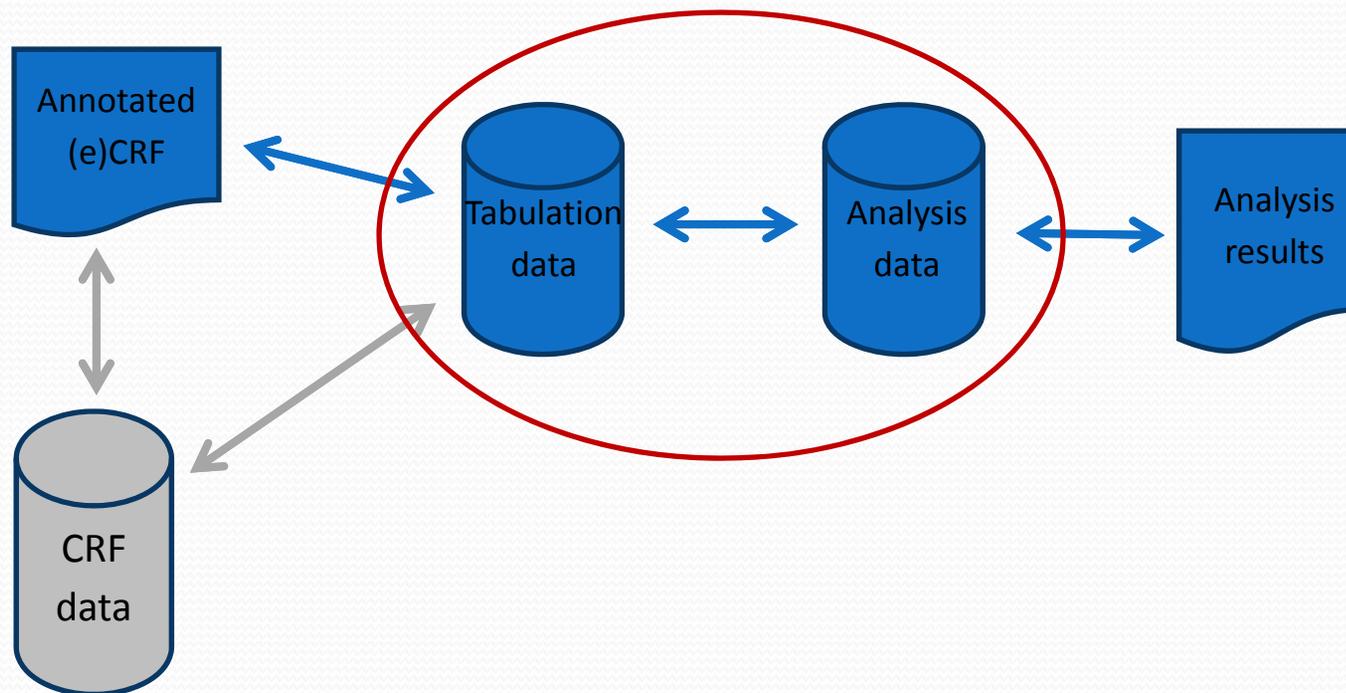
- Two annotated (e)CRFs
 - One for CRF data
 - One for Tabulation data
- Include reference data
 - Row numbers
- Document derivations from CRF to Tabulation data
 - Denote any derived tabulation data
 - Use controlled terminology
 - Minimize transformations
- Create define file and other metadata from specs

Traceability between Data Collection and Tabulation - **CDISC**

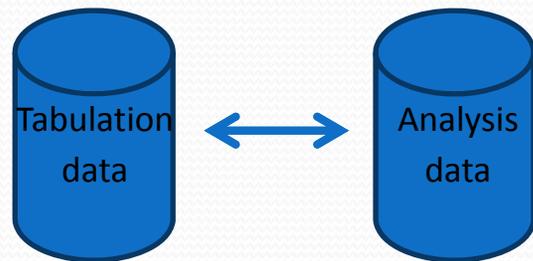


- Two annotated (e)CRFs
 - One for **CDASH** data
 - One for **SDTM** data
- Include reference data
 - Row numbers (**--SPID**)
- Document derivations from **CDASH** to **SDTM** data
 - Denote any derived tabulation data (**--DRVFL**)
 - Use controlled terminology
 - Minimize transformations
- Create define file and other metadata from specs

Traceability in Clinical Trials Data Flow

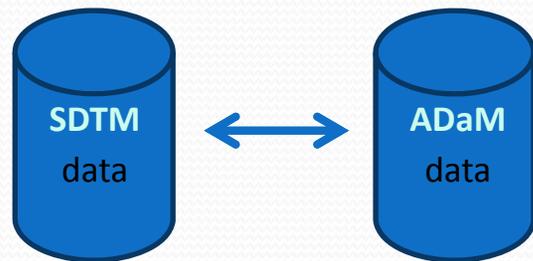


Traceability between Tabulation and Analysis Data



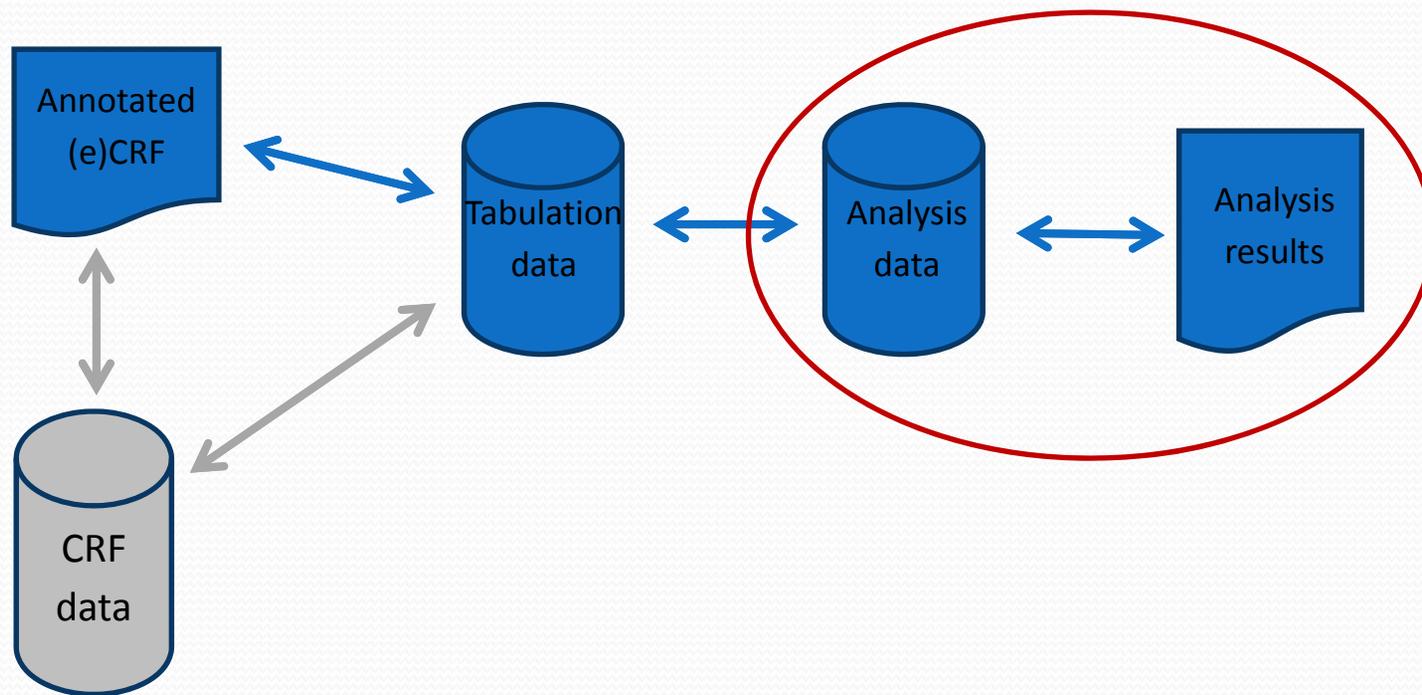
- Use controlled terminology
- Clarify copied vs. derived data
- Use informative labels
- Include intermediate data for complex derivations
- Include reference data
 - Row reference
 - Tabulation data not used for analysis
- Create define file and other metadata from specs

Traceability between Tabulation and Analysis Data - CDISC

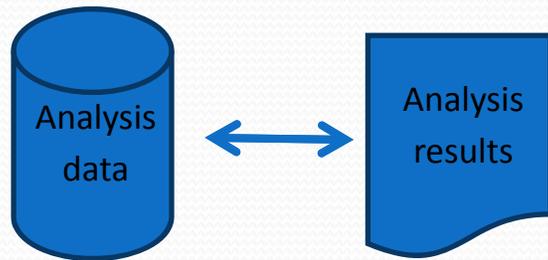


- Use controlled terminology
- Clarify copied vs. derived data (e.g., **DTYPE**, **PARAMTYP**)
- Use informative labels
- Include intermediate data for complex derivations
- Include reference data
 - Row reference (**--SEQ**, **SRCDOM/SRCSEQ**, **ASEQ**)
 - Tabulation data not used for analysis (e.g., ***DTC**)
- Create define file and other metadata from specs

Traceability in Clinical Trials Data Flow

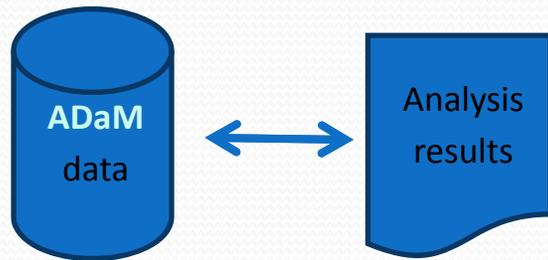


Traceability between Data Analysis and Results



- Ensure analysis datasets are “analysis-ready”
- Use analysis flags to show which rows are included in specific analyses
- Create numeric versions of data that have meaning other than just for sorting
 - Example: numeric visit represents number of weeks

Traceability between Data Analysis and Results - CDISC



- Ensure analysis datasets are “analysis-ready”
- Use analysis flags to show which rows are included in specific analyses (**ANLzzFL**)
- Create numeric versions of data that have meaning other than just for sorting
 - Example: numeric visit represents number of weeks (**AVISITN**)

Metadata

- Metadata components for tabulation and analysis data
 - DEFINE.XML
 - Reviewers Guides
 - SDRG = Study Data Reviewers Guide
 - ADRG = Analysis Data Reviewers Guide
- Most of what we need in metadata already exists
 - Re-use material in specifications, programs, tools

and speaking of
Reviewers Guides...

CSS Wiki Navigation

Navigation trail: [Optimizing the Use of Data Standards](#) » [Traceability and Data Flow](#) » [PhUSE Wiki](#) » [CSS Working Groups](#)

CSS Working Groups

(Redirected from [FDA Working Groups](#))

Instructions

[\[edit\]](#)

This wiki is available to document the progress of the PhUSE Computational Science (CS) working groups and projects. The wiki enables quick, easy and transparent online collaboration. Working groups each have a main page which is the starting page for all Working Group activity.

If you are new to wikis you might want to **check out the help section**. To get started click on your working group below.

For instructions on how to subscribe to the working group/project email list or contact a working group leadership team [click here](#)

The Working Groups are governed by the [Computational Science Steering Committee](#) which provides oversight to the Working Groups.

PhUSE CS Working Groups

[\[edit\]](#)

[CS Dashboard on \[www.phuse.eu\]\(http://www.phuse.eu\)](#)

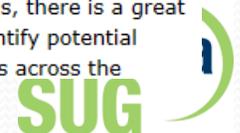
[CS Catalog of Deliverables on \[www.phuse.eu\]\(http://www.phuse.eu\)](#)

[CS Catalog of Deliverables Under Review on \[www.phuse.eu\]\(http://www.phuse.eu\)](#)

Optimizing the Use of Data Standards The development and adoption of data standards over the last decade has shown significant promise in improving efficiencies in the product submission and review process. However, there have also been gaps, issues and challenges in the interpretation and use of the standards. This group will identify specific gaps preventing FDA and industry from optimizing the use of standards and collaborate to close those gaps.

- **NEW:** A new Project within the Optimizing the Use of Data Standards is getting started to develop a templates for a **Study Data Standardization Plan**.. Visit the Wiki page for ~~more information on how you can participate on this new project.~~
- **ADRG and SDRG Work Packages Referenced in Final FDA Study Data Technical Conformance Guide, Dec 2014**

Development of Standard Scripts for Analysis and Programming With the development and implementation of industry data standards, there is a great opportunity to develop standard reporting across industry and to support the needs of FDA medical and statistical reviewers. This working group will identify potential standard scripts for data transformations and analyses across an within therapeutic areas. The goal will be to begin the process of standardizing analyses across the



CSS Wiki Navigation

Completed Projects

Reviewer's Guide Work Packages referenced in FDA Study Data Technical Conformance Guide v2.0, Dec2014

(2014-05-13) Analysis Data Reviewer's Guide (ADRG) - ADaM provides a framework that enables analysis of the data, while at the same time allowing reviewers and other recipients of the data to have a clear understanding of the data's lineage from collection to analysis to results. Although ADaM provides a robust metadata framework, FDA Reviewers benefit from additional, human-readable, documentation of analysis methods, data sets, and programs that cannot be fully explained within the ADaM metadata. The Final ADRG package contains a template to be used in submissions with completion guidelines and examples. The ADRG template provides an orientation to the submitted data in a consistent and usable format.

(2013-05-13) Study Data Reviewers Guide (SDRG) - The define.xml document does not adequately document mapping decisions, sponsor-defined domains, and other key study components and a SDRG would help to address this documentation gap. The goal of this project is to develop a SDRG template jointly between CDER, Industry, and CDISC to be used for submissions.

Version	Release Date	Downloadable Work Package	Changes from Previous Version
v1.1	26-Jan-2015	ADRG Package v1.1 2015-01-26	<ul style="list-style-type: none">• Added ADSL Header to Section 5.2• Improved ADRG Template usability• Minor revisions to instructions in ADRG Completion Guidelines• ADRG Examples updated to match revised SDRG Template
v1.01	13-May-2014	ADRG Package v1.01 2014-05-13	Initial Version
v1.2	26-Jan-2015	SDRG Package v1.2 2015-01-26	<ul style="list-style-type: none">• Removed Trial Design Dataset navigation table from Section 2.3• Improved SDRG Template usability• Minor revisions to instructions in SDRG Completion Guidelines• SDRG Examples updated to match revised SDRG Template
v1.1	03-May-2013	SDRG Package v1.1 2013-05-13	Initial Version

SDRG Package Contents

This package contains the following files:

[SDRG_Template_2015-01-26.doc](#)

This file is the template you would start with to produce an SDRG.

[SDRG_Completion_Guidelines_2015-01-26.pdf](#)

This document explains how to use the template to build your SDRG document.

Under the Samples folder you will find the following additional files:

[study-data-reviewers-guide-example-001a-2014-08-06.doc](#)

[study-data-reviewers-guide-example-001a-2014-08-06.pdf](#)

[study-data-reviewers-guide-example-002a-2014-08-06.doc](#)

[study-data-reviewers-guide-example-002a-2014-08-06.pdf](#)

[study-data-reviewers-guide-example-abc123-2014-08-06.doc](#)

[study-data-reviewers-guide-example-abc123-2014-08-06.pdf](#)

These sample Study Data Reviewer's Guide documents are included to illustrate the intended use of the template and demonstrate what a finished document should look like. Both the Word doc version and pdf version are included so you can see how the Word documents were constructed and also how they are rendered in pdf format for submission.

SDRG Package Updates

Changes in Version 1.2

The following changes were made to the SDRG_Template document based on user feedback about Version 1.1:

1. The Trial Design Datasets navigation table was dropped from section 2.3.
2. The header style for the Trial Design Dataset descriptions (sections 2.3.x) was changed from automatic numbering to manual numbering because the automatic numbering was easily broken and difficult to fix.
3. The automatic numbering of the dataset description headers in section 3.3 was similarly changed to manual numbering for the same reasons.

The SDRG_Completion_Guidelines document was updated to align some of the examples with the changed template document, clarify some practices and update the instructions for numbering headers in sections 2.3 and 3.3 of the template.

All of the sample SDRG documents were updated to reflect the changed template and instructions.

ADRG Package Contents

This package contains the following files:

[ADRG_Template_2015-01-26.doc](#)

This file is the template you would start with to produce an ADRG.

[ADRG_Completion_Guidelines_2015-01-26.pdf](#)

This document explains how to use the template to build your ADRG document.

Under the Samples folder you will find the following additional files:

[analysis-data-reviewers-guide-sample-9999-2014-10-16.doc](#)

[analysis-data-reviewers-guide-sample-9999-2014-10-16.pdf](#)

[analysis-data-reviewers-guide-sample-cdisc-pilot-2014-10-16.doc](#)

[analysis-data-reviewers-guide-sample-cdisc-pilot-2014-10-16.pdf](#)

These sample Analysis Data Reviewer's Guide documents are included to illustrate the intended use of the template and demonstrate what a finished document should look like. Both the Word doc version and pdf version are included so you can see how the Word documents were constructed and also how they are rendered in pdf format for submission.

ADRG Package Updates

Changes in Version 1.1

The following changes were made to the ADRG_Template document based on user feedback about Version 1.0:

1. The style of the Section 5.2 dataset sub-headers was changed to use manual numbering rather than automatic numbering, because the automatic numbering would sometimes break.
2. The numbering styles in optional sections and subsections (ones that can be deleted) were changed to prevent the numbering from breaking when other numbered sections are deleted.
3. The formatting of the section 5.2 dataset index table was updated (font, underlining and point size) and an ADSL row was added with a hyperlink.
4. An ADSL sub-header placeholder was added as section 5.2.1
5. Tables were given a consistent look and formatting.

The ADRG_Completion_Guidelines document was updated to align the examples and instructions with the changed template document, clarify some practices and add more specifics to the pdf production instructions.

All of the sample ADRG documents were updated to reflect the changed template and instructions.

Why Do Reviewers Need Guides?

- Define.xml is good at describing variable- and value-level details, but doesn't have a place for higher level information about the study conduct and population, datasets, complex derivations involving multiple variables, etc.
- If a derivation description requires formatting (tables, equations, multiple paragraphs) it has to go in a document. Define.xml does not support formatting.
- The Technical Conformance Guide says include them.

Show and Tell

- Study Data Reviewer's Guide
Template and Guidelines
- Analysis Data Reviewer's Guide
Template and Guidelines

Usage Tips for the Templates

- Keep every section in the finished Guide unless the Guideline says you can delete it. It's easier for reviewers when the document format is standard!
If a section does not apply to your study, mark it Not Applicable.
- Some sections have manual numbering (trust me, it's better that way) so double check that things end up numbered correctly.
- Remember to update the TOC before you pdf it.
- Don't try to "fix" formatting. No good will come of it.

Not Just for SDTM and ADaM

- The ADRG is designed to document analysis issues generally, not ADaM data specifically. You should be able to use it pretty much as-is for any analysis data.
- The SDRG has some SDTM-specific content. You can remove those sections or table columns when using the template for “Item 11” data.

References

- CSS wiki: phusewiki.org -> CSS Working Groups
 - White paper on Linear Data Flow: -> Optimizing the Use of Data Standards
 - ADRG and SDRG: -> CSS Catalog of Deliverables
- CDISC: cdisc.org

Questions?