

## Standard Analyses and Code Sharing Working Group Update

Nancy Brucken, Clinical Solutions Group, Inc.;  
Dante Di Tommaso, Omeros;  
Jane Marrer, Merck;  
Mary Nilsson, Eli Lilly and Company;  
Jared Slain, Charles River Laboratories  
Hanming Tu, Frontage Labs

### ABSTRACT

This poster updates the community on the efforts of the six project teams in the PHUSE Standard Analyses and Code Sharing Working Group. The Working Group publishes recommended analyses of clinical data suitable across therapeutic areas. These publications include presentations (tables, listings and figures) of the results from those analyses. The Working Group's GitHub repository contains a wealth of scripts that have been written by PHUSE members, or developed and contributed by the FDA and other organizations. The collaborative efforts of this group improve our collective efforts to design and implement transparent and robust analyses of our clinical data for regulatory decision making. Leveraging crowd-sourcing code development of these recommended analyses can promote access to and adoption of these analyses, and lead to better data interpretations and increased efficiency in our clinical drug development and review processes.

### INTRODUCTION

The PHUSE Standard Analyses and Code Sharing Working Group consists of six project teams that focus on different aspects of the analysis and reporting of clinical trials data that is common across therapeutic areas. These teams have been in existence for several years now, and have produced quite a few deliverables that are available for download from the PHUSE website, including white papers recommending certain types of analyses and graphical displays of safety data, as well as a GitHub repository of standard scripts to produce those displays. This work was last presented at PharmaSUG in 2014, so this paper is intended to serve as an update and report on recent deliverables, as well as those currently in progress from each of the project teams.

The goals of the Working Group are as follows:

1. Establish and maintain a publicly available repository for storing program code to be used as analytical tools for medical research.
2. Where gaps exist, develop recommendations for analyses and displays in areas that could benefit from crowd-sourcing.
3. Where gaps exist, develop code for recommended analyses and displays that could benefit from crowd-sourcing (to reside in the repository).

PHUSE Working Groups depend on contributions from volunteers. Project teams actively recruit new members to help with these efforts. We hope that this paper supports the work and inspires volunteers to participate.

### PROJECT TEAM UPDATES

Each of the six project teams comprising the Standard Analysis and Code Sharing Working Group meets regularly. This paper will provide an update on the activities of each of the project teams.

### ANALYSIS AND DISPLAY WHITE PAPERS

CDISC standards have been established to cover trial designs (PRM), data collection systems (CDASH), storage of observed data (SDTM), and analysis dataset structures (ADaM). However, no standards have as yet been created to cover data displays- the tables, listings and figures created to show the results of each clinical study. The purpose of the Analysis and Display White Papers project team is to develop

white papers that provide recommended tables, figures and listings of commonly-collected data for use in clinical trial study reports and submission documents. Six white papers have been finalized since the inception of this project team, as noted in Figure 1.

The project team conducted a Safety Analytics Workshop at the June 2019 Computational Science Symposium, which was recorded, and can be downloaded and viewed in segments at <https://www.phuse.eu/white-papers>. Types of data addressed in the workshop included adverse events, clinical laboratory results, subject disposition and concomitant medication.

The project team is also planning to release several white papers for public review in 2020. Here is a complete schedule showing all past and upcoming white paper releases:

	Version 1		Version 2	
	Review	Publish (ed)	Review	Publish (ed)
Vital Signs, Labs, ECG – Central Tendency		Oct 2013 ✓		
Non-Compartmental PK		March 2014 ✓		
Demographics, Disposition, Medications		Oct 2014 ✓	Q4 2017 ✓	Q1 2018 ✓
Vital Signs, Labs, ECG – Outliers / Shifts		Sept 2015 ✓		
QT/QTc Studies		March 2016 ✓		
Adverse Events		Feb 2017 ✓		
General Output Tips and Considerations (Karin)	Q1 2020	Q2 2020		
Treatment-Emergent Definitions (Survey Results)	Q2 2020	Q3 2020		
Labs (Wei)	Q3 2020	Q1 2021		
Hepatotoxicity (Terry)	Q3 2020	Q1 2021		
Questionnaires (Karin)	Q3 2020	Q1 2021		
Listings	Q3 2020	Q4 2020		
Safety Topics of Interest (Brenda)	Q3 2020	Q4 2020		
Vital Signs	Q3 2021	Q4 2021		
ECG's	Q3 2021	Q4 2021		
Treatment-Emergent Definitions (Recommendations on Definition)	Q3 2021	Q1 2022		
Interactive Displays for Clinical Safety Data (Wei)	Q4 2021	Q1 2022		

**Figure 1. Past and Future White Papers**

### CODE SHARING (REPOSITORY)

The purpose of the Code Sharing project team is to establish and maintain a collaboration platform leveraging crowd-sourcing to improve content and implementation of analyses recommended by the Analysis and Display White Papers project team. Incorporation of these analysis will lead to improved data interpretations and increased efficiency in the clinical drug development and review process. In order to achieve this goal, the project team is collaboratively building shared tools and a reusable code library.

Since the phuse-scripts repository was created in Github in 2013, the team has recommended coding style and guidelines, and developed a qualification process to review, develop and share the scripts. Many scripts have been contributed by the FDA and volunteers, and are hosted in the GitHub repository (<https://github.com/phuse-org/phuse-scripts>), including Spotfire templates and R and SAS programs to produce tables and figures from the 2013 Vital Signs, Labs and ECG white paper.

The project team has also released 2 white papers, one showing the [displays produced by FDA-contributed scripts](#), and the other describing [the script metadata required by the repository](#). It also recently developed an [R Shiny tool for creating the simplified TS domain](#) required for submission of nonclinical data using SEND.

## **TEST DATA FACTORY**

The Test Data Factory project team is working to provide SDTM- and ADaM-conformant test data to support systematic and comprehensive testing of applications developed for processing and analysis of clinical study data. It had previously released a version of the second CDISC pilot ADaM datasets that could be used with the script repository to produce tables and figures from the Analysis and Display White Papers. In 2018 and 2019, both the pilot SDTM and ADaM datasets were updated to conform to SDTMIG v3.2 and ADaMIG v1.1, respectively. These datasets are freely available, and can be downloaded from <https://www.phuse.eu/white-papers>.

The project team is now developing R and SAS code to create test datasets based on user input. The team are progressing as follows:

1. Define the scope of test data – datasets, user inputs, characteristics, etc.
2. Establish an interface for user requirements for a test database, and conventions for configuring data elements. For example, setting the study population size, treatments, and treatment allocation required to simulate a study database
3. Define and implement SAS and R scripts to generate the user-specified test datasets.
4. Bonus step: Provide an infrastructure that enables users to easily use scripts to create test datasets.

As of the end of 2019, the TDF team has designed a basic framework for specifying and simulating an SDTM database, and initiated implementation of the SDTM.DM domain based on user requirements.

## **COMMUNICATION, PROMOTION AND EDUCATION**

The Communication, Promotion and Education project team conceptualizes efficient ways to communicate the Working Group's progress and results. It relies on the acceptance, input, feedback and further development from/by the user community. It also defines target groups, timing, communication channels and the means for presenting this information in order to ensure the success of the Working Group.

A major focus of 2019 was the development of an educational website to assist with navigating the GitHub repository. This site was developed in Squarespace, and launched in the Third Quarter of 2019 at <https://education.phuse.eu/coderepository>.

Plans for 2020 include various conferences where the project team will present updates, as well as PHUSE monthly webinars and other channels. Additionally, work will begin on an educational site centered around Safety Analytics that will collate the knowledge base developed by the Analysis and Display White Papers project.

## **BEST PRACTICES FOR QUALITY CONTROL AND VALIDATION**

There are no industry standards or guidelines on the best way to validate and QC programs used for analysis and reporting. Pharmaceutical companies and CROs follow quality control processes to identify errors in their analysis programs and ensure high quality. However, the processes used vary throughout the industry.

The Best Practices for Quality Control and Validation project team is authoring a white paper describing the best approach for this process. The white paper was published in April 2020, and available at <https://www.phusewiki.org/docs/WorkingGroups/Deliverables/Best%20Practices%20for%20Quality%20Control%20and%20Validation-%20White%20Paper%20.pdf>.

## GOOD PROGRAMMING PRACTICES (GPP) IN MACRO DEVELOPMENT

Macros provide an effective way to automate and reuse code in a standard and consistent manner across SAS programs. The ability to reuse code means that the use of GPP is particularly important in macro code. There is a need to develop a consensus and document good programming practices specifically for macro development, covering such areas as coding style, structured documentation, refactoring, optimization and saving compiling time, and best practices for writing. The purpose of the project team is to develop guidelines for creating well-structured and precisely documented macro code that will be easy to read and maintain over time.

## CONCLUSION

The PHUSE Standard Analysis and Code Sharing Working Group, consisting of six separate project teams, has been actively working towards its goal of providing recommendations for the analysis and display of commonly-collected clinical trials data, as well as test data and tools that can be used for the development of those analyses and displays. The Working Group is also focusing on best practices for macro/script creation in general, as well as for validation/QC of those programs.

All of the project teams are actively recruiting volunteers to help with these efforts, except for the Best Practices for Quality Control and Validation team, which is nearly finished with its white paper. If you are interested in helping out, please contact [wendy@phuse.eu](mailto:wendy@phuse.eu).

## REFERENCES

PHUSE. "Final White Papers & References from Working Group Projects". Accessed January 3, 2020. Available at <https://www.phuse.eu/white-papers>.

## ACKNOWLEDGMENTS

Many thanks to Jane Marrer for creating the slides on which this paper is based, and to the Standard Analysis and Code Sharing Working Group project team leads for their review of the paper.

## CONTACT INFORMATION

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

Nancy Brucken  
Clinical Solutions Group, Inc.  
[nbrucken@csg-inc.com](mailto:nbrucken@csg-inc.com)

Dante Di Tommaso  
Omeros  
[dditommaso@omeros.com](mailto:dditommaso@omeros.com)

Jane Marrer  
Merck  
[jane\\_marrer@merck.com](mailto:jane_marrer@merck.com)

Mary Nilsson  
Eli Lilly and Company  
[nilsson\\_mary\\_e@lilly.com](mailto:nilsson_mary_e@lilly.com)

Jared Slain  
Charles River Laboratories  
[jared.slain@crl.com](mailto:jared.slain@crl.com)

Hanming Tu  
Frontage Labs  
[htu@frontagelab.com](mailto:htu@frontagelab.com)

Any brand and product names are trademarks of their respective companies.