

PharmaSUG 2023 - Paper DS-148**Automation of the SDSP CBER Appendix for Vaccine Studies**

Nicole Jones, Merck & Co., Inc., Rahway, NJ, USA
Pritesh Solanki, Merck & Co., Inc., Rahway, NJ, USA
Jagruti Patel, Merck & Co., Inc., Rahway, NJ, USA

ABSTRACT

The Study Data Standardization Plan (SDSP) is an important means used to communicate between the sponsor and FDA. The SDSP describes the data standards used by the sponsor in non-clinical and clinical studies across the drug development program. Vaccine studies that are submitted to the FDA Center of Biologics Evaluation and Research (CBER) require an additional appendix. Part of this appendix includes 4 sections of which two are SDTM sections: (i) list of all SDTM datasets with all the expected or permissible variables used/planned for each study in the analysis and (ii) list of all SUPPQUAL variables used/planned for use. Manually generating these tables could take weeks if no data are available. This manual process is also prone to errors in the form of missing variables/domains. In this paper, we present a Shiny application developed to semi-automate the generation of these tables. The Shiny application can utilize SDTM mapping specs or SDTM datasets as input for generating the two appendix sections. Although some manual review is still required, this tool greatly reduces the time for table generation from weeks to days and improves accuracy.

INTRODUCTION

The use of a Study Data Standardization Plan (SDSP) was first introduced in the July 2016 FDA Study Data Technical Conformance Guide (TCG) v3.1. In that TCG, including the SDSP is identified as a non-binding recommendation for sponsors to plan, prepare and share their company position on the use of data standards prior to a regulatory submission. The early interaction promotes a two-way communication for setting expectations implemented during product development.

With the SDSP template available from PhUSE, and the FDA welcoming earlier sponsor interaction, sponsors now have an opportunity to present the use of data standards throughout the product development lifecycle beginning with non-clinical data in the IND stages and through the NDA\BLA and post-marketing commitments. Advance sharing of the SDSP enables sponsors to drive decisions on legacy data conversions or plan for up-versioning of data that would allow for data integration across multiple studies.

CBER SDSP CHECKLIST

Through the PhUSE and FDA collaboration, the SDSP template was streamlined to support both Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) submissions. In March 2018, PhUSE released the SDSP (v1) template that incorporated a CBER Appendix that included additional background on the planned data standards. This paper highlights the application developed to streamline and automate some of the steps in authoring the SDSP CBER Appendix. We also present authoring considerations while addressing challenges, advantages, and timing of regulatory interaction to support a CDER or CBER SDSP review.

When an SDSP is prepared for CBER, authors should allocate more time to address the CBER Appendix. In the initial PhUSE template (released March 2018), the SDSP template included a CBER Appendix that contained five additional sections to facilitate CBER review. These sections include: Introduction, SDTM Datasets, Supplemental Qualifiers, ADaM Datasets, and Integrated Summary of Safety and Efficacy (ISS and ISE).

Unlike the SDSP body, which contains an inventory of all non-clinical and clinical trials and the trial status (completed, ongoing, planned), the CBER Appendix contains the additional information from clinical trials that are planned for the actual submission. In other words, the CBER Appendix is a subset of the trials listed in the SDSP body. The next five sub-sections describe the authoring considerations for creating the CBER Appendix.

SDTM Datasets

For each trial included in the IND, CBER requests a table that lists all planned SDTM datasets. The table is similar to the SDTM Datasets section found in the tabulations section of Clinical Study Data Reviewer's Guide (CSDRG). However, unlike the CSDRG that is shared at the time of the Biologic License Agreement (BLA), the CBER Appendix - SDTM Datasets section is intended to be shared during the trial development phase.

The section identifies a list of planned domains, regardless of whether data are present. Organized by domain class (trial design, special purpose, interventions, events, findings, findings about, Relationships and custom), the SDTM Datasets section includes information about the SDTM domains, variables used for analysis that are not required by the SDTM model and custom domains used by the sponsor. This section provides background to the CBER reviewers so that they understand the purpose of the data. Information presented in the SDTM Datasets section could result in CBER making recommendations including but not limited to remapping data to avoid custom domains.

Supplemental Qualifiers

CBER is seeking a list of supplemental qualifiers planned for each clinical trial. The Supplemental Qualifier section includes a comprehensive list of supplemental domains (SUPPXX), qualifier variable name and label (QNAM, QLABEL) and origin as it relates to a corresponding question on the CRF or derivation information.

Data are often mapped to supplemental qualifiers more than necessary. CBER has expressed that supplemental variable are difficult to mine given a sponsor's approach to placement of data that does not fit the SDTM model. During the review, sponsors may find CBER offering suggestions to remap data that would display data in supplemental qualifiers.

When developing the CBER Appendix, sponsors should identify the supplemental qualifiers contributing to the key analysis. Relevant information identifying the variable's purpose will help avoid additional feedback.

ADaM Datasets

The ADaM Datasets section is included for each clinical trial with a planned analysis. Organized by data structure (subject level analysis dataset, basic data structure, occurrence data structure), the ADaM Datasets table identifies the name of the analysis dataset and its labels.

ISS and ISE

In comparison to an SDSP body which lists the trials included in the pooled analysis, CBER recommends sponsors to identify the potential pooled/integrated analysis (efficacy, safety etc.). The additional information found in the CBER Appendix should also identify the clinical trial phase and contributing datasets used to support the analysis.

Although the information would eventually be shared in the Analysis Data Reviewer's Guide (ADRG), the early identification promotes early planning and collaboration for the expected analysis.

CBER APPENDIX GENERATOR

The CBER Appendix Generator currently supports creation of the SDTM datasets section and Supplemental Qualifier section. As mentioned above, these two sections contain the list of planned SDTM domains (regardless of the presence of data) and a complete list of SDTM supplemental qualifiers. Generating these two sections can be challenging given data is not always present for all domains and all supplemental qualifiers. This lack of data meant that authoring of these sections relied on a manual process where programmers had to review each individual SDTM mapping specification file, compile a list of variables by domain, then remove all the required variables keeping only those which are expected and permissible. This manual process takes weeks and is prone to errors.

To address the challenges faced when generating the SDTM Datasets and Supplemental Qualifiers sections of the CBER Appendix, we leveraged R and Shiny to develop an application to streamline and semi-automate the process. Shiny is an R package that makes it easy to build interactive web apps straight

from R. By leveraging shiny, we created an intuitive user interface (UI) to help streamline the process of generating the SDSP appendix tables. The application utilizes SDTM mapping specs or SDTM datasets as an input. The application then removes required SDTM variables and generates a list of only expected and permissible variables. Finally, the application generates the SDTM table (appropriately grouping variables by domain and domain class) and the SUPPQUAL table. Users are able to further specify the SDTM standard version being used (SDTM3.2 vs SDTM3.3) to ensure the appropriate variables are included for their study.

CBER APPENDIX GENERATOR USER INTERFACE

Figure 3. Figure 1 shows the initial UI once the application is opened. Users are prompted to select either (1) “Initial SDSP” or (2) “Subsequent/Final SDSP”. When selecting the first option, SDTM mapping specs will be used for table generation. When selecting the second option, SDTM datasets are used for table generation. Users are then prompted to select the SDTM version being used for their study. The tool is currently designed to support SDTM versions 3.2 and 3.3.


[CBER Appendix Generator](#)
[Home](#)
[Instructions](#)
[Configuration](#)

Please select an option from the drop down menu

Please select SDTM version for your study

mkcberappendix 0.1.0 | Developer & Maintainer: Nicole Jones (BARDS) | Hosted on RStudio Connect

Figure 1. Initial UI

After users make their initial selections, additional input fields and button are displayed. Figure 2 shows the subsequent screen. Users are able to select the appropriate study folders and SDTM mapping spec folder before clicking the “Generate SDSP Table” button. The application will then read either the mapping specs or SDTM datasets, process and display the resulting tables in the UI. A metadata file is maintained external to the tool. It contains the role of the variables and the class of each dataset.

Please select an option from the drop down menu

Initial SDSP

Please select SDTM version for your study

SDTM 3.2

Compound Indication (CPI Folder)*

SDTM Folder*

Study Spec Folder

Generate SDSP Table

mckberappendix 0.1.0 | Developer & Maintainer: Nicole Jones (BARDS) | Hosted on RStudio Connect

Figure 2. UI after initial user input

Figure 3 shows the UI with the preview of SDTM table. The tables are displayed to allow users to review what has been generated and review for any major concerns (i.e. missing domains or variables). If any mapping specs were not able to be processed, the UI will display a list of such files. Users are expected to review these files manually for any variables that may need to be added to the table manually. Both the SDTM table and Supplemental Qualifiers table can be exported to .rtf files. After generating these files, manual intervention is still needed. Submission consultants, SDTM, and Analysis & Reporting programmers are expected to collaborate to identify the appropriate variables to keep in the table.

[Home](#)
[Instructions](#)
[Configuration](#)

Please select an option from the drop down menu

Initial SDSP

Please select SDTM version for your study

SDTM 3.2

Compound Indication (CPI Folder)*

mk9999

SDTM Folder*

prot001

Study Spec Folder

sdm32_specs

Generate SDSP Table

Download SDTM Domains and Variables

Download SUPPQUAL

Show 10 entries

Generated for mk9999-001

Search:

Domain	Select Domains to be Submitted	Variables to be utilized (besides required)	Class	Additional Comments
1	AE (Adverse Events)	<X>	EVENTS	
2	DS (Disposition)	<X>	EVENTS	

Figure 3. User Interface after Table Generation

CONSISTENCY OF SDTM MAPPING SPECIFICATIONS

As mentioned above, the application will not always be able to process the mapping specifications. For the application to read the mapping specifications, the files must have the same standard structure. Given that the mapping specifications are manually maintained, there is the possibility that some of the specs may deviate from the standard structure. When the tool encounters mapping specs that do not conform to the expected structure, a list is generated and displayed in the UI to inform users of the issue. Users can then take two actions: (1) investigate and resolve the structure issue or (2) review the specs and add any missing variables to the tables generated by the application.

In addition to structure issues, file format can also be an issue. Our mapping specifications are saved as both .xlsx files or .csv files. When saving the files as .csv the encoding used for the file is important. The application would encounter an issue if the file was saved with an .csv extension but encoded as Unicode Text (.txt).

Figure 4 shows an example of this issue. Although the file is saved as Book1.csv, the format is Unicode Text. This encoding leads to the standard csv reading functions in R to fail. Users again have options for action: (1) manually review these files and add any missing variables to the tables or (2) save the files with the field being updated to .csv and regenerate the tables using the tool.

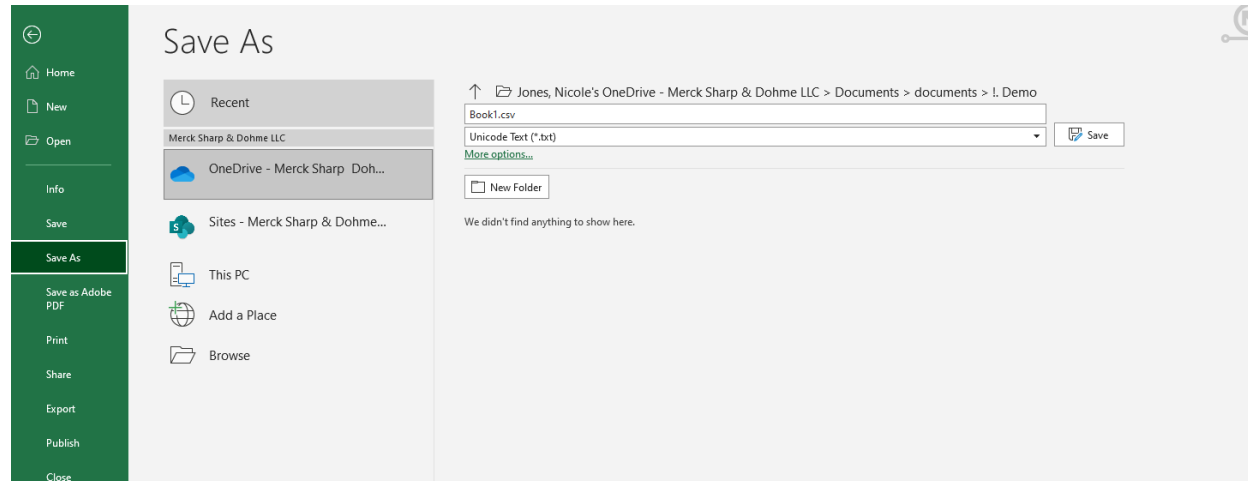


Figure 4. CSV Encoding Issue

CONCLUSION

The information presented in this paper offer authors suggestions and considerations while creating the SDSP body and CBER Appendix. An application to streamline creation of the CBER Appendix is presented as well. Drawing from Merck experience while implementing the SDSP, it required cross functional collaboration to identify key roles to contribute to the SDSP content. Even with defined roles and responsibilities, authors need to reserve ample time to compile the content and avoid underestimating the work ahead when addressing products with a large volume of trials, CROs and partnerships, and complex programs. Accurate and timely sharing of information presented in the SDSP helps drive data standard decisions to support the product development.

The development of tools such as this CBER Appendix generator can help to streamline the process and reduce the time spent on generating the CBER Appendix. Prior to the introduction of the CBER Appendix generator, the tables were manually generated which led to typos, mismatching variables and SDTM domains, and often took weeks to generate. Since implementing the use of this application in our internal workflow, production of the CBER Appendix section has been reduced from weeks to days, there has been an increased clarity around what should be included in the CBER Appendix, and the accuracy of the tables being generated has increased. Overall, our programmers have reported the great benefits this application has provided for CBER Appendix generation workflow.

In addition, sufficient details in the CBER Appendix that describe the purpose and content of custom datasets and supplemental qualifiers will provide CBER reviewers an early understanding of a sponsor's use of data. The advance sharing of the SDSP and CBER Appendix at the recommended prior to end of Phase 2 promotes early regulatory interactions that set expectations for an efficient and meaningful review.

Although the SDSP template only became available in March 2018, authors of this paper plan to share lessons learnt on the use of SDSP and CBER Appendix back to the Optimizing Data Standards Working Group. The next step is to offer suggestions to the PhUSE Working Group so that the SDSP Completion Guideline and SDSP Sponsor Implementation Guide will incorporate additional considerations for authoring the SDSP and CBER Appendix.

REFERENCES

PhUSE Study Data Standardization Plan Template (v1)

<https://www.phuse.eu/documents//sop/wp/phuse-tp001-study-data-standardization-plan-v1-8409.docx>

PhUSE Study Data Standardization Plan Completion Guideline (v1)

<https://www.phuse.eu/documents//sop/wp/phuse-wp001-sdsp-completion-guideline-v1-8407.pdf>

ACKNOWLEDGMENTS

The authors would like to recognize the following subject matter experts from our statistical programming departments who partnered with us to develop this application: Helen Walters, Donna Hyatt, and Suhas Sanjee.

CONTACT INFORMATION

Nicole Jones
Merck & Co., Inc.
126 E Lincoln Ave
Rahway, NJ 07065
(w) (732) 5942249
nicole.jones3@merck.com

Pritesh Solanki
Merck & Co., Inc.
351 N. Summneytown Pike
North Wales, PA 19454
(w) 267-305-8224
pritesh_solanki@merck.com

Jagruti Patel
Merck & Co., Inc.
351 N. Summneytown Pike
North Wales, PA 19454
(w) (215) 6315122
jagruti.patel3@merck.com