PHUSE US CSS 2020 : PP09

SEND Data Factory

Bob Friedman, Xybion; Kevin Snyder, FDA; Eli Miller, Atorus Research; Daniel Russo, FDA

Abstract

onuse

The PHUSE non-clinical scripts team is developing a SEND Data Factory tool which can create example SEND datasets. This will allow industry stakeholders to produce example SEND datasets for use in the testing visualizations, data exchange and tools development.

The program is in a preliminary but working state, written in R-Shiny and shared on the PHUSE Github for group collaboration or downloading and customization. This tools allows for the selection of the study design (number of dose groups and animals per group), desired measurement (body weights, clinical observations, micropathology, etc.) and it then can output .xpt dataset files with synthetic data.

SEND data factory	× +								-	-	Ċ
← → C ③ 127.0.0.1:4	573					Q 🕁	Ø	0	G	x	
SEND data factory	=										
Output settings	<										
🛗 Study design 💉	SEND IO	G structure	5								
	Domain	Column	Туре	Label	Codelist	Expectance	у				
Study Name to create:	DM	STUDYID	Char	Study Identifier		Req					
StudyABC	DM	DOMAIN	Char	Domain Abbreviation	DM	Req					
	DM	USUBJID	Char	Unique Subject Identifier		Req					
Test article:	DM	SUBJID	Char	Subject Identifier		Req					
TestA123	DM	RFSTDTC	Char	Subject Reference Start Date/Time	ISO 8601	Req					
	DM	RFENDTC	Char	Subject Reference End Date/Time	ISO 8601	Ехр					
Select Study Type:	DM	RFXSTDTC	Char	Date/Time of First Study Treatment	ISO 8601	Perm					
O Single-dose	DM	RFXENDTC	Char	Date/Time of Last Study Treatment	ISO 8601	Perm					
Multi-dose	DM	SITEID	Char	Study Site Identifier		Perm					
Carcinogenicity	DM	BRTHDTC	Char	Date/Time of Birth	ISO 8601	Perm					
 Safety Pharm - Respiratory Safety Pharm - Cardiovascular 	DM	AGE	Num	Age		Perm					
 Early Fetal Development 	DM	AGETXT	Char	Age Range	number-number	Perm					
	DM	AGEU		Age Unit	(AGEU)	Exp					
Select Treatment Groups:	DM	SEX	Char	Sex	(SEX)	Reg					
2 Control group	DM	SPECIES		Species		Perm					
Group 2: Low dose			Char		(SPECIES)						
 Group 3: Mid dose Group 4: High dose 	DM	STRAIN	Char	Strain/Substrain	(STRAIN)	Perm					
	DM	SBSTRAIN	Char	Strain/Substrain Details		Perm					
Select set options:	DM	ARMCD	Char	Planned Arm Code		Exp					
🛛 TKanimals	DM	ARM	Char	Description of Planned Arm		Perm					
Recovery animals	DM	SETCD	Char	Set Code		Req					
M. Animal information	CO	STUDYID	Char	Study Identifier		Req					
Animal information	° co	DOMAIN	Char	Domain Abbreviation	CO	Req					
A Data selections	< <u>co</u>	RDOMAIN	Char	Related Domain Abbreviation		Ехр					
	со	USUBJID	Char	Unique Subject Identifier		Exp					
» Produce Data	< co	000UD	Char			Danna					

Future work is planned to improve the realism of the data by allowing configuration of species and age specific ranges, simulated growth and dose responses, and the desired distribution of pathological observations.

Methods:

- This tool is written in R-Shiny for GUI and data manipulation.
- Numerical data is synthesized through random data within configurable ranges by specie. Finding observations data is synthesized through random data based upon configurable expected frequency of occurrence.
- A successful dataset will be defined as one that passes a validator with no errors and that has data that is scientifically possible.

Figure 1: Study selections (left) and domain structure

Show structure		dditional Save TSPARMC 1 AGETXT 2 AGEU	trial summary (edita D TSPARM Age Text	ble) TSVAL					
 Study design Addt'l trial summary 		Save TSPARMO 1 AGETXT	D TSPARM						
 Study design Addt'l trial summary 		TSPARMO 1 AGETXT		TSVAL					
Addt'l trial summary		TSPARMO 1 AGETXT		TSVAL					
Addt'l trial summary		1 AGETXT		TSVAL					
			Age Text						
		2 AGEU		10-50					
Dosing configuration			Age Unit	DAYS					
Dosing configuration		3 DOSDUR	Dosing Duration	P29D					
		4 EXPENDT		2019-07-05	2019-07-05				
		5 EXPSTDT		2019-08-03					
Animal information	~	6 GLPTYP	Good Laboratory Practice Type	GLP					
Select species:			P Housing Type	Group housed					
		8 ROUTE	Route of Administration	INTRAVENOUS1					
BOVINE		9 SDESIGN	Study Design	PARALLEL	PARALLEL				
		10 SPLRNAM	Test Subject Supplier	Supplier B	Supplier B				
		11 SPREFID	Sponsor's Reference ID	S1234567	S1234567				
elect strain:		12 SSPONSO	R Sponsoring Organization	The sponsor	The sponsor				
129/SV -		13 STCAT	Study Category	Toxicity					
123/34					-				
		14 STDIR	Study Director	John Doe					
elect sex:		15 STITLE	Study Title	The Study Title					
		16 🕆 👘 > This	PC > Downloads > MyStudy (62).zip		~ C	Search MyStudy (62).zip			
M		17 🔲 Name	Туре	Compressed size Password Size	Ratio	Date modified			
🖞 F		18 🗋 bw.xpt	SAS Xport Transport File	4 KB No	30 KB 88%	6/8/2020 2:03 PM			
		19 dm.xpt	SAS Xport Transport File	1 KB No	5 KB 86%	6/8/2020 2:03 PM			
		ds.xpt	SAS Xport Transport File	1 KB No	5 KB 88%	6/8/2020 2:03 PM			
elect animals Per Group:		20 ex.xpt	SAS Xport Transport File	1 KB No	6 KB 90%	6/8/2020 2:03 PM			
) 4		21 mi.xpt	SAS Xport Transport File SAS Xport Transport File	3 KB No 1 KB No	55 KB 95% 9 KB 90%				
		22 ta.xpt	SAS Xport Transport File	1 KB No	4 KB 82%				
8		te vot	SAS Xport Transport File	1 KB No	3 KB 79%				
16		23 ts.xpt	SAS Xport Transport File	2 KB No	5 KB 75%	6/8/2020 2:03 PM			
		🗋 tx.xpt	SAS Xport Transport File	TKB No	6 KB 86%	6/8/2020 2:03 PM			
20	•								

Use cases:

- Demonstrating data visualization tools
- Testing systems that receive data from the evolving standards
- Creating large data sets for testing data loading
- As example code to show:
 - R script creation of XPT files
 - R script reading of the CDISC SENDIG standard (currently PDF)
- R script reading of NCI Control Terminology files (xls version)

Features

- Controlled Terminology version selection
- Study type, species selections
- Number of groups and animals per group
- Recovery and TK groups
- Trial summary data and dosing settings







Figure 2: Trial summary entry and resultant dataset download

SEND data factory																	
Output settings	< Dataset	s Created	ł														
 Study design Addt'l trial summary 	< TS	BW CL	ма мі	LB	DS D	DM TX	TA TE										
Dosing configuration	STUDYID		USUBJID	FOCID	MISEQ	MIGRPID	MIREFID	MISPID	MITESTCD	MITEST	MIBODSYS	MIORRES	MISTRESC	MIRESCAT	MICHRON	MIDISTR	MISTAT
Animal information	< MyStudy MyStudy		MyStudy- 1 MyStudy-	NA	1	NA	NA	NA	MIEXAM	NA	NA	UNERMARKABLE	UNERMARKABLE	NA	NA	NA	NA
 Data selections Produce Data 	< MyStudy		1 MyStudy-	NA	3	NA	NA	NA	MIEXAM	NA	NA	INFLAMATION	INFLAMATION	NA	NA	NA	NA
Produce datasets	MyStudy	мі	1 MyStudy- 1	NA	4	NA	NA	NA	MIEXAM	NA	NA	INFLAMATION	INFLAMATION	NA	NA	NA	NA
Show datasets	MyStudy	мі	MyStudy- 1	NA	5	NA	NA	NA	MIEXAM	NA	NA	UNERMARKABLE	UNERMARKABLE	NA	NA	NA	NA
Download Data	MyStudy	МІ	MyStudy- 1	NA	6	NA	NA	NA	MIEXAM	NA	NA	UNERMARKABLE	UNERMARKABLE	NA	NA	NA	NA
	MyStudy		MyStudy- 1	NA	7	NA	NA	NA	MIEXAM	NA	NA	HYPERPLASIA	HYPERPLASIA	NA	NA	NA	NA
	MyStudy	MI	MyStudy- 1 MyStudy-	NA NA	8	NA	NA	NA	MIEXAM	NA	NA	INFLAMATION	INFLAMATION	NA	NA	NA	NA
	MyStudy		1 MyStudy-	NA	10	NA	NA	NA	MIEXAM	NA	NA	UNERMARKABLE	UNERMARKABLE	NA	NA	NA	NA
	MyStudy	мі	1 MyStudy-	NA	11	NA	NA	NA	MIEXAM	NA	NA	HYPERPLASIA	HYPERPLASIA	NA	NA	NA	NA
	MyStudy	мі	4 MyStudy-	NA	12	NA	NA	NA	MIEXAM	NA	NA	UNERMARKABLE	UNERMARKABLE	NA	NA	NA	NA

Figure 3: Dataset tabbed view

Next steps:

- Ability to create custom and experimental domains
- Simulated time progress for numerical endpoints
- Simulated dose responses for endpoints
- Create test datasets with validation rule violations

Conclusion:

This tools can be used by an industry group or an organization to create proof-of-concept datasets. The tools can be used in the future to test changes and enhancements to SEND IG by



Figure 4: Nonclinical Script **Figure 5:** SEND Data Factory Figure 6: Public website Assessment PhUSE Wiki page **Github Project** running SEND Data Factory

giving stakeholders an application to interact with changes in a practical way.

Note: The opinions expressed in this poster are those of the authors and do not necessarily represent the opinions of their respective organizations.