

AN UPDATE ON CDISC-SEND DATA SUBMISSIONS TO FDA

Jesse Anderson, MPH
Lead, Nonclinical Program

Christopher Bartlett, PhD
Contractor, IBM

Office of Computational Science
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

PHUSE US Connect 2020
March 9, 2020

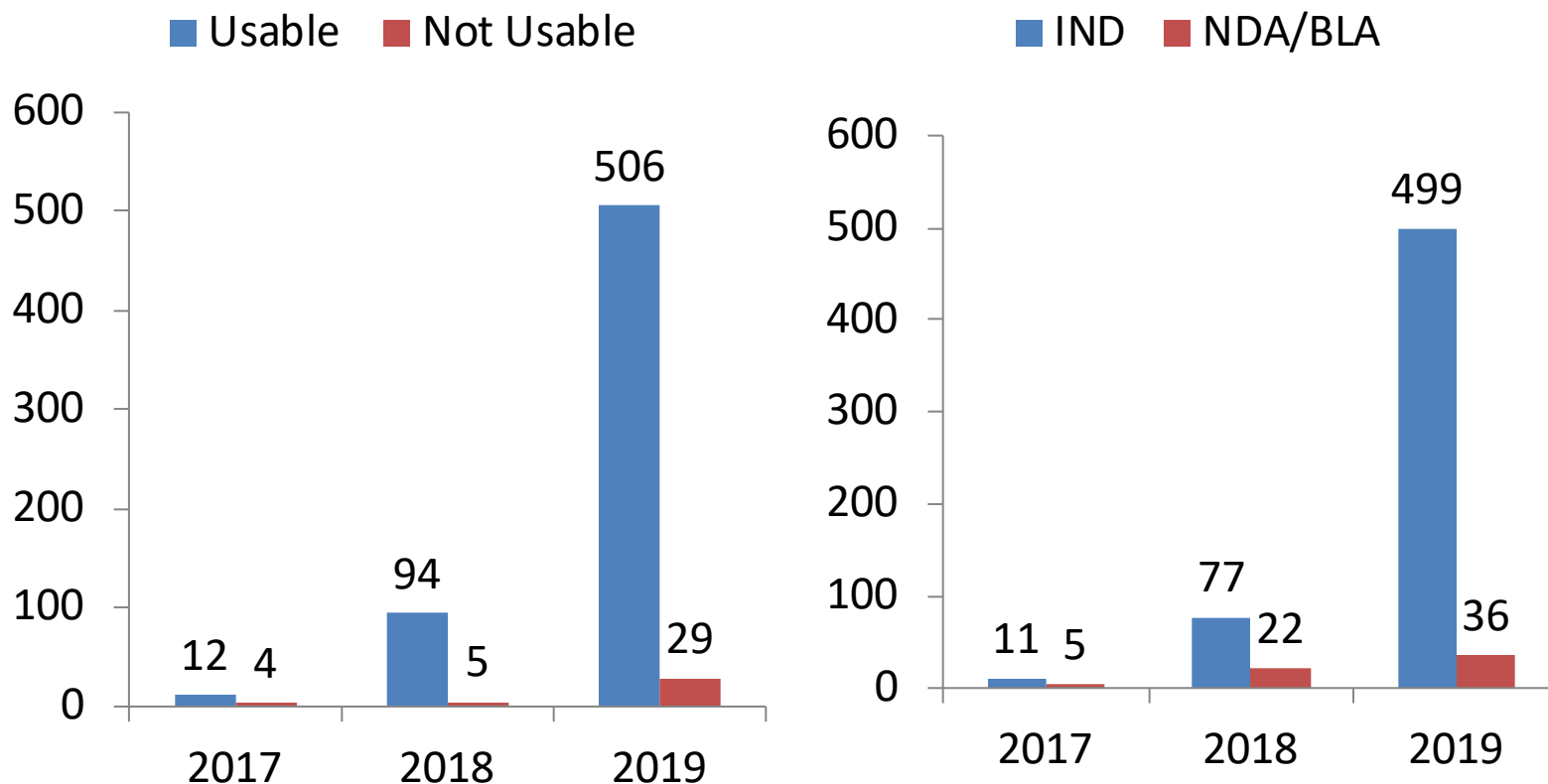
The views and opinions presented here represent those of the speakers and should not be considered to represent advice or guidance on behalf of the U.S. Food and Drug Administration.

FDA Submissions Requirements



- Standardized electronic data required for certain submissions to FDA
- Change in process for nonclinical reviewers
- Quick timeline for many nonclinical submissions
- Needed a way to introduce standardized electronic data to reviewers and use these data as part of review

Submission of Standardized Nonclinical Study Data Continues to Increase



Number of Applications Per Calendar Year

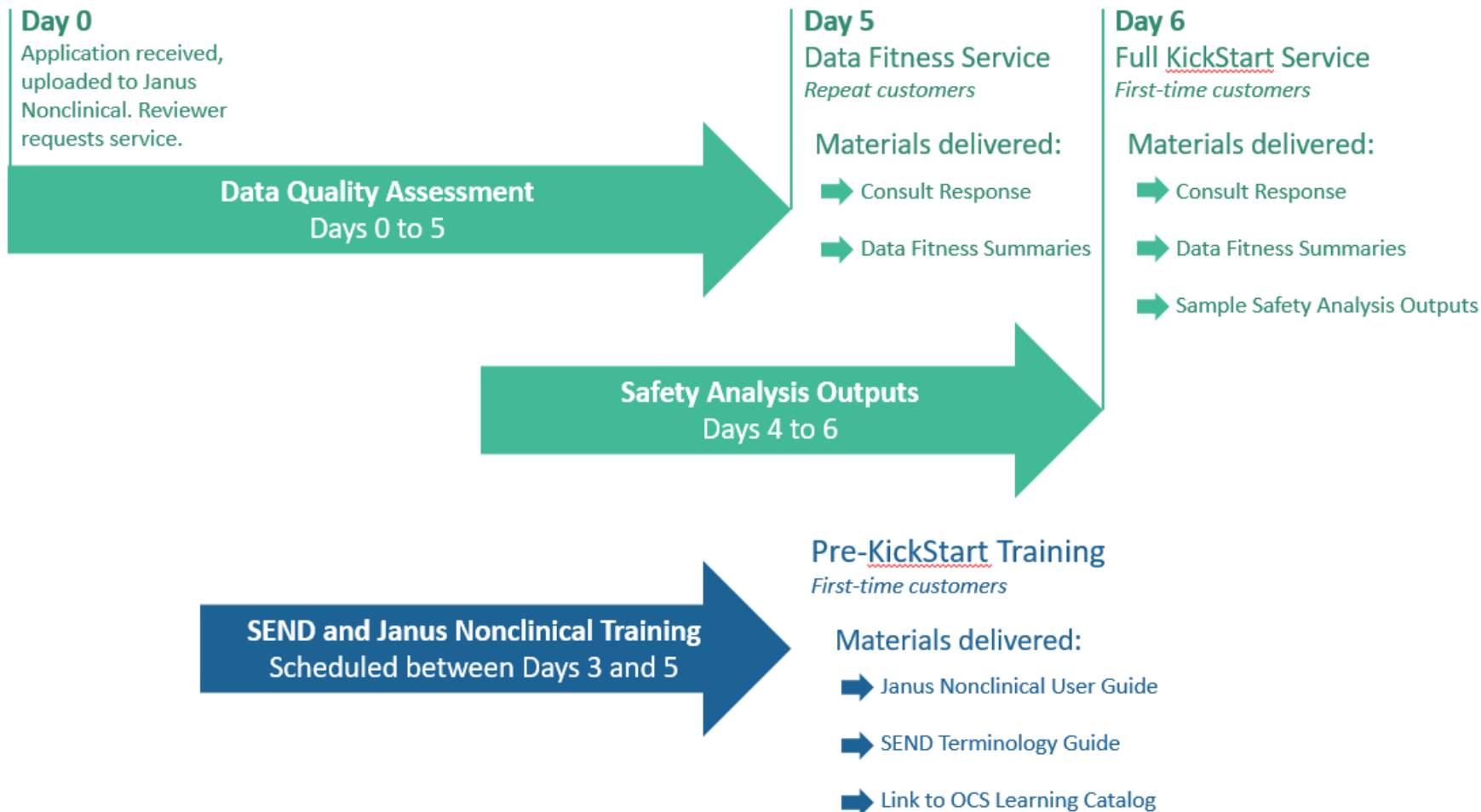
The types of studies included Repeat-Dose Toxicity, Single-Dose Toxicity, and Carcinogenicity.
FDA has received 650 applications with SEND data during this time period.

KickStart Service



- Provides reviewers with an overview of SEND concepts and the systems used for reviewing SEND datasets
- Delivers a data quality assessment to show issues with the SEND datasets in an application
- Trains reviewers on data exploration in Janus Nonclinical, a proprietary system for FDA review

KickStart Timeline Overview



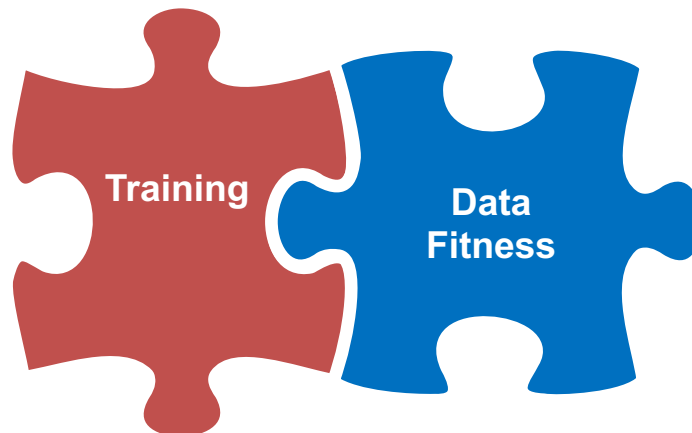
Pre-KickStart Training

- For first KickStart service, the KickStart team provides reviewers with general training on the SEND topics, including:
 - Domains
 - Controlled Terminology
 - Overview of the nSDRG
 - Introduction to the define file
 - Introduction to Janus Nonclinical



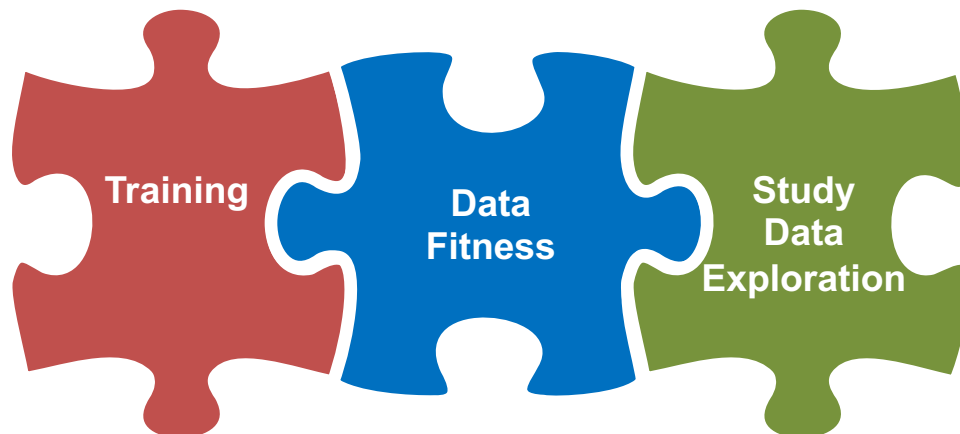
KickStart Data Fitness Assessment

- Automated and manual review of SEND datasets and associated nSDRG and define.xml files
 - Verify compliance with standards and FDA rules/recommendations
 - Confirm and document data not submitted
 - Check consistency across study files and documents
 - Ensure summarizations included in study report can be reproduced
- Issues that affect data analysis are discussed with reviewer
- Sponsor data fitness report details all issues identified



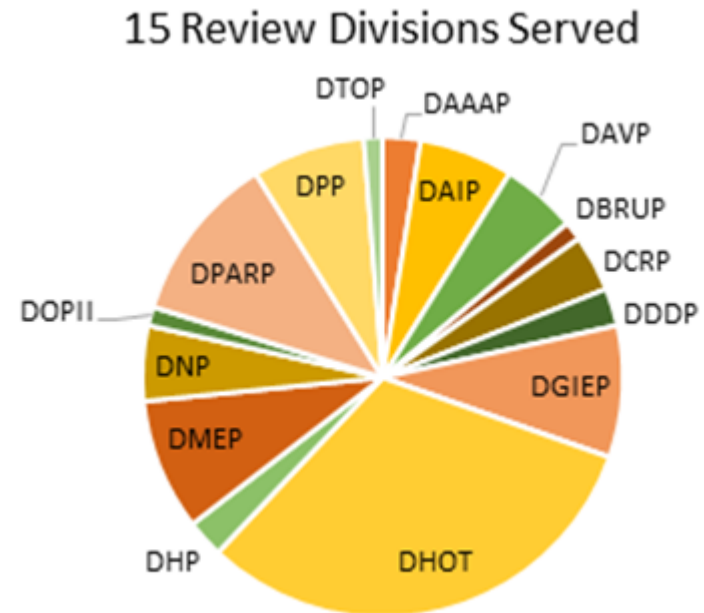
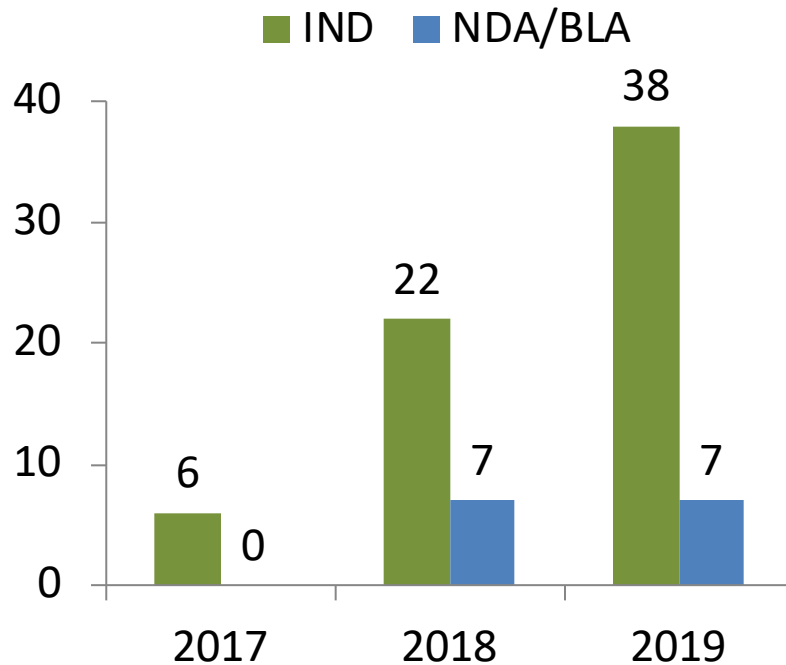
KickStart Data Exploration Session

- Interactive look at the best way to interact with application study data using Janus Nonclinical
- Show tables and graphs from key domains with findings aligned with study report when possible
- Provide outputs may be used as part of application review documents



KickStart Services for Nonclinical Review

Applications Through End of December 2019



Calendar Year Counts

Kickstart service provided when:

- Study loaded into Janus Nonclinical
- Reviewer requests service

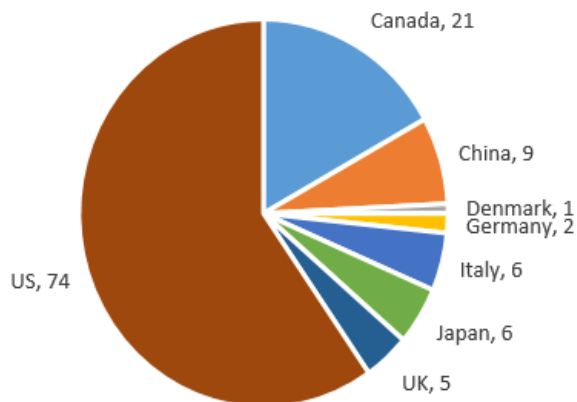
One or two studies per application generally reviewed

Studies Included in Kickstart Applications

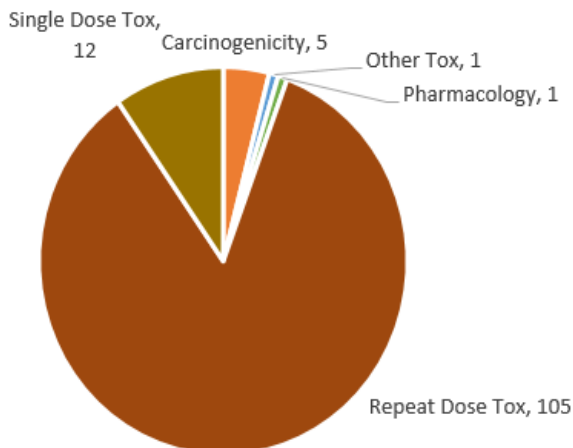
Applications Through End of December 2019



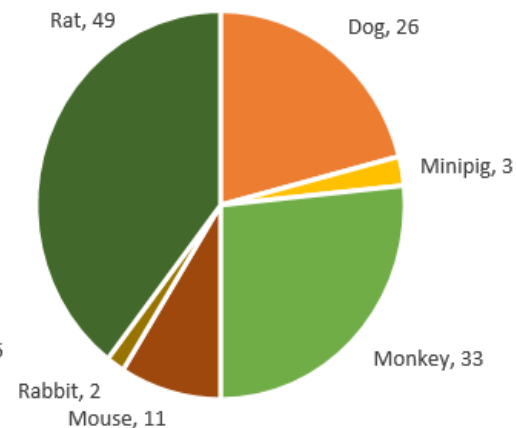
Test Facility Countries



Study Type



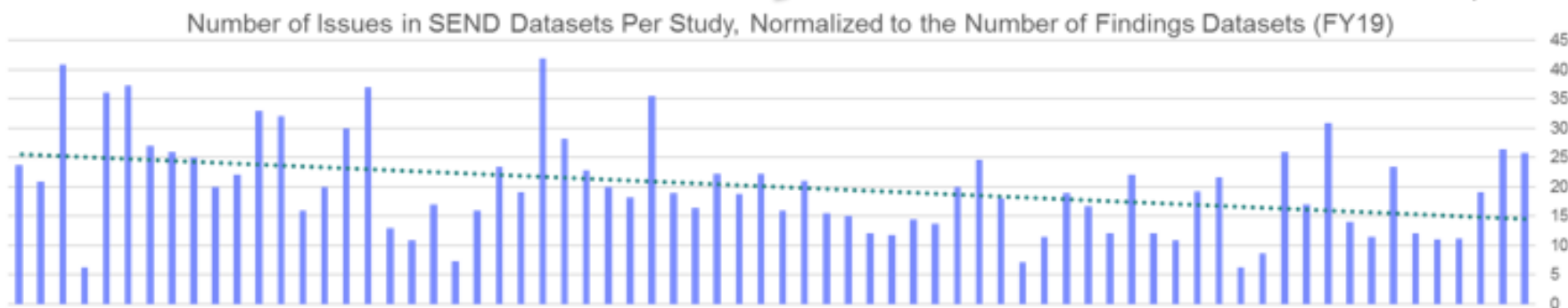
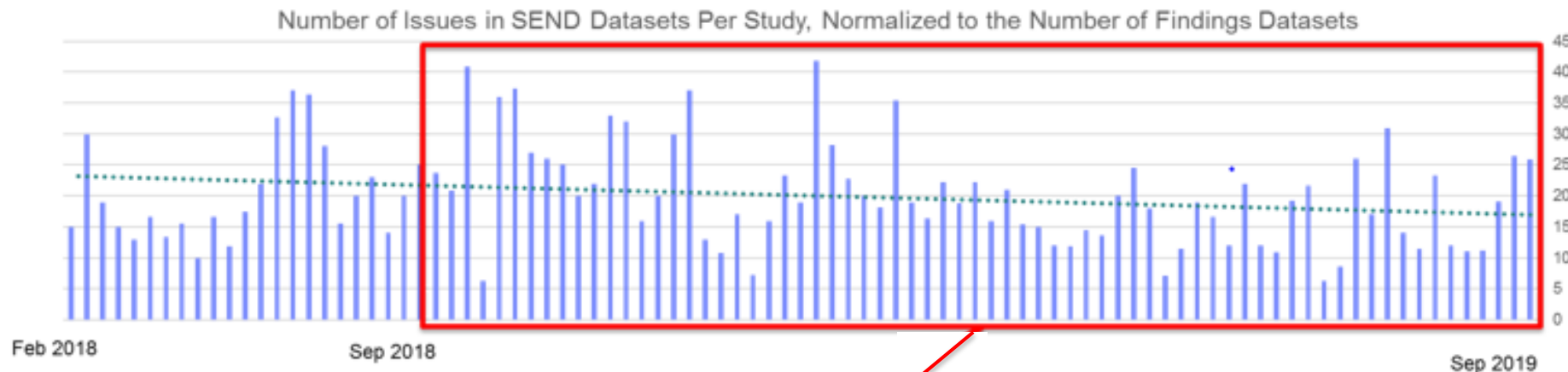
Species



Quality Improvement in KickStart Studies



Trend Line Moving in the Right Direction!



Common Data Fitness Issues

Analysis of 70 studies in 47 applications for fiscal 2019 identified a number of key themes:

- Incorrect reporting of timing variables needed for summarization and analysis of results
- Incorrect reporting of categorical data
- Omission of the numeric value to use in calculations as a replacement for text result
- Undefined codes/abbreviations included in datasets
- Issues with alignment of organ/tissue names with the study report

Less Common but Serious Issues

- Incorrectly named files, for example:
 - pool.xpt; bw-body-weights.xpt, data-definitions.xml
- Incorrect formatting of ISO8601 values
 - Duration variables most problematic
- Text in an ISO8601 variable
 - Includes TSVAL when TSPARMCD requires date or duration
- Multiple group labels (TXPARMCD=GRPLBL) associated with a single group code (TXPARMCD=SPGRPCD)
- Animals assigned to the wrong set

Common Issues with Standardization of Timing Variables

- Planned day does not align data with the study report summary
- Incorrect reporting of timing for tests or observations scheduled relative to dose
- Incorrect reporting of result from unscheduled tests or observations.

VISITDY/--NOMDY Used for Analysis

Study Schedule indicates each animal has two sample collections: One pretest and one at the conclusion of dosing. Sponsor reported actual collection day in VISITDY.

Reviewer must adjust days prior to use of data in summary analyses.

GROUPED-SUMMARY : NUMERIC

GROUPED SUMMARY : NON-NUMERIC

DAILY BY ANIMAL

ANIMAL DETAILS

Filter table data by term...

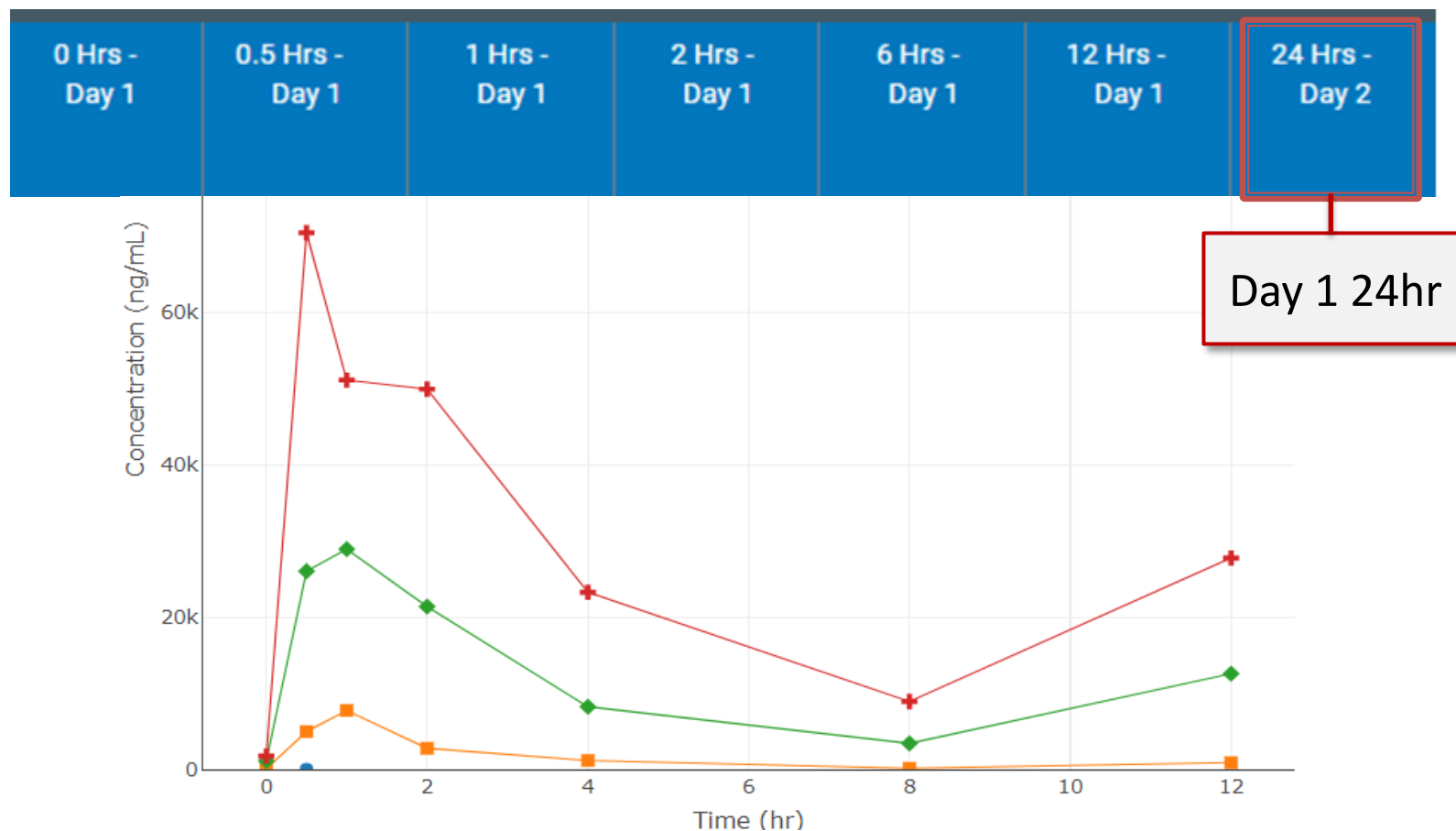
Group	S...	Test	Animal ID	Unit of the Standa... Result	Day -13	Day -12	Day -11	Day 178	Day 179
☐ 1-Control ...	Male	Alanine Aminotransferase	1	U/L			21		17
☐ 1-Control ...	Male	Alanine Aminotransferase	2	U/L		34		26	
☐ 1-Control ...	Male	Alanine Aminotransferase	3	U/L	47			30	
☐ 1-Control ...	Male	Alanine Aminotransferase	4	U/L	23			30	
☐ 1-Control ...	Male	Alanine Aminotransferase	5	U/L			29		30
☐ 1-Control ...	Male	Alanine Aminotransferase	6	U/L			25		27
☐ 1-Control ...	Male	Alanine Aminotransferase	7	U/L		32			21
☐ 1-Control ...	Male	Alanine Aminotransferase	8	U/L	21			22	
☐ 1-Control ...	Male	Alanine Aminotransferase	9	U/L	18			14	

Pretest

Post Dose

VISITDY/--NOMDY and –ELTM Use for Analysis

Sponsor reported sample collection day in VISITDY when samples collected at times relative to dose. Tabular data is therefore mislabeled; graphs of results relative to dose do not include data for timepoints > 24 hrs post dose



Timing Relative to Dose

- Elapsed Time --ELTM:
 - uses ISO8601 duration format
 - describe amount of time relative to the dose event described in --TPTREF and --RFTDTC, on day in VISITDY/--NOMDY
- For timed infusion (or other dosing over time):
 - Is --ELTM relative to start or end of dose?
 - Does --TPTREF describe this?
- For multiple doses per day:
 - Which dose is --ELTM relative to?
 - Does --TPTREF describe this?

From the SENDIG: It is recommended that --TPTREF be as descriptive as possible so the reference time point can be inferred without looking at other variables.

Unscheduled Tests/Observations

- Results from unscheduled tests or observations are generally tabulated separately in the study report tables
- In SEND, for an unscheduled test result or observation:
 - SENDIG 3.0, VISITDY with no value identifies an unscheduled result
 - In SENDIG 3.1, --USCHFL=Y” identifies an unscheduled result
- Correct flagging of unscheduled results in SEND allows reviewer to align results with study report
 - Unscheduled data can be included/excluded
 - Using only –TPT to label unscheduled results is not sufficient

VISITDY	LBDTC	LBDY	LBTPT	LBTPTNUM
15	2018-01-23	15	Unsch Vet Request	6
22	2018-01-30	22		•
22	2018-01-30	22		•
22	2018-01-30	22		•

Data Fitness Issues with Reporting of Categorical Test Results

- Seen in the Laboratory Test Results (LB) dataset
- Issue with including LBSTRESN for categorical results
 - Refer to the FDA Technical Conformance Guide, section 4.1.3.3
 - SEND has no specific flag for categorical results
 - Define-xml value-level metadata inconsistently reported
 - Janus attempts to identify non-numeric tests using LBSTRESN

Example Incidence Counts for LB in Janus



GROUPED-SUMMARY : NUMERIC

GROUPED SUMMARY : NON-NUMERIC

DAILY BY ANIMAL

ANIMAL DETAILS

Filter table data by term...

Category	Test	Day	Value	Unit of the Standardiz... Result	Male				Female			
					Contol [Termina...	2 mg/kg [Terminal,...	6 mg/kg [Termi...	20 mg/kg [Terminal,...	Contol [Termi...	2 mg/kg [Termi...	6 mg/kg [Termina...	20 mg/kg [Termin...
Urine Sediment	Erythrocytes	-1	N	/HPF	1	1	1		1		1	
Urine Sediment	Erythrocytes	-1	0-1	/HPF					1		1	
Urine Sediment	Erythrocytes	-1	1-3	/HPF	1	1						
Urine Sediment	Erythrocytes	-1	>100	/HPF			1					
Urine Sediment	Erythrocytes	90	N	/HPF		1			1		1	1
Urine Sediment	Erythrocytes	90	0-1	/HPF								1
Urine Sediment	Erythrocytes	90	1-3	/HPF					1			
Urine Sediment	Erythrocytes	90	3-5	/HPF							1	
Urine Sediment	Erythrocytes	90	5-10	/HPF		1						
Urine Sediment	Erythrocytes	120	N	/HPF							1	
Urine Sediment	Erythrocytes	120	3-5	/HPF							1	

Tests with no LBSTRESN reported with summary incidence counts in Janus Nonclinical

Example of Incorrect Standardization



Day	Category	Test	Unit of the Stand... Result		Male			
					Group 1,0mg/kg/d... [Terminal,R...	Group 4,600-400mg/kg/... [Terminal,R...	Group 2,60mg/kg/... - Terminal	Group 3,200mg/k... - Terminal
☰ ☐ -8								
	URINALYSIS	Protein	g/L	Mean	0.000 (Text)	0.000 (Text)	0.000	0.000 (Text)
	URINALYSIS	Protein	g/L	Std Dev	±0.000 (Text)	±0.000 (Text)	±0.000	±0.000 (Text)

Animal ID	Day	Result	Result in Character
1001	-8	0 0	
1002	-8	0 0	
1003	-8	0 0	
1004	-8	0.3-0.7	
005	-8	0 0	
1006	-8	0.1-0.2	

CLOSE
DAILY BY ANIMAL
ANIMAL DETAILS

Semi-quantitative urine protein test are in SEND with the result range in LBSTRESC and with values incorrectly in LBSTRESN. This causes Janus Nonclinical to report means and standard deviations on those numeric results rather than incidence counts.

Data Fitness Issues with Character Replacement Values

- Sponsor omitted character replacement values for use in PC or LB summary calculations in approximately half the studies reviewed:
 - Most commonly missing for plasma concentration results reported in the PC dataset.
- FDA Tools cannot replicate and use group mean results supplied in study report when replacement values are not submitted.

Example: LBCALCN Not Supplied



Day	Specimen Material Type	Sub Category	T Test	Unit of the Stan... Result		Male				Female			
						0 mg/kg	1 mg/kg	3 mg/kg	10 mg/kg	0 mg/kg	1 mg/kg	3 mg/kg	10 mg/kg
29	PLASMA	Clinical Chemistry	Sorbitol Dehy...	U/L	Mean	1.500 (Text)							
	PLASMA	Clinical Chemistry	Sorbitol Dehy...	U/L	Std Dev	±0.577 (Text)							
	PLASMA	Clinical Chemistry	Sorbitol Dehy...		N	12 (Text)							
58	PLASMA	Clinical Chemistry	Sorbitol Dehy...	U/L	Mean	2.200 (Text)							
	PLASMA	Clinical Chemistry	Sorbitol Dehy...	U/L	Std Dev	±0.837 (Text)							
	PLASMA	Clinical Chemistry	Sorbitol Dehy...		N	6 (Text)							

Animal ID	Day	Result	Result in Character
10001	29	2	2
10002	29		<1.0
10003	29	1	1
10004	29		<1.0
10005	29	1	1
10006	29		<1.0
10007	29		<1.0
10008	29	2	2
10009	29		<1.0
10010	29		<1.0
10011	29		<1.0

U/L	
2.0	
0.5 }	
1.0	
0.5 }	
1.0	
0.5 }	
0.5 }	
2.0	
0.5 }	
0.5 }	
0.5 }	
0.5 }	
n	12
Mean	0.83
S.D.	0.58

} : Below the LLOQ, represented as a half value of LLOQ

Study Report indicates that half of LLOQ is used for results <1.0. LBCALCN not supplied so results <1.0 eliminated from mean in FDA Tools

Undefined Code/Abbreviations included in the Dataset



- Almost 50% of studies have codes/abbreviations in the SEND datasets in variables where they are not expected
 - Frequently without a definition
 - For some studies, the nSDRG notes generally that abbreviations are defined in the study report.
- Codes and abbreviations have been seen as result values, in reason fields, as finding modifiers and in comments.
- Undefined unit abbreviations (not part of the published UNIT codelist) have also been submitted.

Examples of Undefined Code/Abbreviations



Test	Sign	Location of a Finding	Comments
Clinical Observation	scaly skin	dose site	D

Organ/Tissue	Test	Exclusion Flag	Reason for Exclusion
GLAND, THYROID/GLAND, PARATHYROID	Weight	Y	E

Category	Test Code	Test	Standardized Result in Character Format
URINALYSIS/URINE CHEMISTRY	COLOR	Color	C0

Category	Test Code	Test	Standardized Result in Character Format	Unit of the Standardized Result
URINALYSIS	CREATVOL	Creatinine Volume	115.4	umol/cp
URINALYSIS	GLUCVOL	Glucose Volume	9.5	umol/cp

Issues with alignment of organ/tissue names with the study report



- Representation of the organs and tissues from the study report tissue list are frequently simplified in SEND datasets
 - Causes ambiguity
 - Can result in invalid summarization
 - Limits use for cross-study analysis

Issues with alignment of organ/tissue names with the study report



- Examples
 - One of a paired tissue weighed but OMLAT not filled
 - The study report indicates that thyroid and parathyroid glands are weighed together, but OMSPEC contains only “GLAND, THYROID”.
 - The study report indicates that “Muscle (biceps femoris)” is required to be examined microscopically, but MISPEC contains only “MUSCLE, SKELETAL” with region unidentified.
 - Multiple regions of the spinal cord are in the list of tissues to examine; MISPEC contains multiple records per animal with MISPEC=“SPINAL CORD” with no information in MISPEC or MIANTREG to identify the region associated with each entry.

Key Points



- FDA saw significant increase in number of studies containing SEND datasets in the past year
- FDA Pharm/Tox Reviewers transitioning to utilize SEND datasets alongside the study report
- Complete and correct SEND datasets are critical for seamless, confident use of SEND datasets by FDA reviewers
- Some common issues in a SEND dataset can complicate or even prevent FDA reviewers use of those SEND datasets
- The FDA KickStart team identifies issues in SEND data to communicate with industry, identify trends, and help reviewers maximize use of their SEND datasets

