



**European Annual Conference:
Nonclinical Topics**



Supporting Documents for SEND Submissions: Nonclinical Study Data Reviewers Guide (nSDRG)

20th June 2017

PhUSE European Computational Science Symposium
London, United Kingdom

[Steven Polley, GSK, UK – on behalf of nSDRG PhUSE working group]



1. Introduction
 - putting the nSDRG into perspective
2. Brief Overview of Content
 - with pointers
3. Feedback from FDA via Fit-For-Use Pilot
 - personal experience with a 'submitted' nSDRG, and general FDA feedback
4. Summary



Introduction:

Why do we need a nSDRG?



From the [Study Data Technical Conformance Guide](#):

Current version 3.3 March 2017

“The preparation of relevant (Study Data) Reviewer Guides is recommended as an integral part of a standards-compliant study data submission.”

Hence – the ‘nonclinical Study Data Reviewers Guide’ (nSDRG)

nSDRG Created & Maintained by the [nSDRG PhUSE Working group](#)



Introduction:

Why is the nSDRG integral?



- The FDA reviewer has the Study Report (with the tabulated data)
- They have the SEND data sets (and define file)
- Why is the nSDRG required? – Why is it *integral*?
 - Essentially the nSDRG serves a number of important purposes but the main objective is to help the reviewer *understand* the Data
 - It should describe how the SEND data relates to the study report
 - To enable the reviewer to make sense of the data
- It is split into a number of sections to achieve this.....



1. Introduction
2. Study Design
3. Standards, Formats, Terminologies, and their Versions
4. Description of Study Datasets
5. Data Standards Validation Rules, Versions, and Conformance Issues
6. Description of Sponsor decisions Related to Data Standards Implementations

[Example used is from the nSDRG submitted as part of the 'Fit-For-Use' Pilot with FDA (see later)]



Overview of Content:

1. Introduction



- Provide study ID information to link document with SEND data sets
- Describe to Reviewer SEND data set creation process
- Indicate to Reviewer a confirmation that SEND data sets accurately represent the study report# and if needed, guide where differences are described.

NB: FDA could ask for proof [retain documented QC Check]

1.1 Study Title, Number, and Version

Study Title	Compound: 4-Week Oral Toxicity Study in Non-Rodents
Study Number	123456
Study Version	There have been no report amendments.

1.2 Summary of SEND Dataset Creation Process

All in-life, clinical pathology, and post-mortem data were collected using Provantis LIMS (Instem). Electrocardiograph measurements were measured and recorded using Easy ViewECG™ and reported in Excel. Plasma drug concentrations were determined by LC-MS/MS analysis and collected using SMS2000 (GSK In-house system); WinNonLin (PharSight [part of Certara]) was used to calculate pharmacokinetic parameters. All pharmacokinetic measurements were processed to excel for provisioning. Input from the each of the collection systems and excel files was processed by the Provantis Centrus SEND solution (Instem) via specific data adaptors to produce an integrated SEND dataset, a define.xml and PDF file, and a validation report. The Provantis Centrus SEND solution and the LIMS-specific adaptors are Part 11 compliant.

1.3 SEND Dataset Verification

Data in the SEND datasets are an accurate representation of the data for Study No. 123456. Any differences between the data sets and the report are described in section 6.2. Verification procedures and documentation supporting this are available upon request.



Overview of Content:

2. Study Design



- Provide a brief text and visual representation of the SEND Trial Design to help the Reviewer relate the report study design to the SEND Trial Design datasets

The following diagram illustrates the trial design.

Study Group	Trial Arms		Element in each Epoch		Trial Set	
	ARMCD	ARM	Screening	Treatment	SETCD	SET
1	1	Control	Screening	0mg/kg/day	1	Control
2	2	Low	Screening	0.3mg/kg/day	2	Low
3	3	Intermediate	Screening	3mg/kg/day	3	Intermediate
4	4	High	Screening	30mg/kg/day	4	High

Domain
Overview



Overview of Content:

3. Standards, Formats, Terminologies & their Versions



- Provide tabular overview of all standards used in the study data package [important!]
- Explain the standards and versions selection

3.1 Standards Used

Standard or Dictionary	Standard or Dictionary	Versions Used
Tabulation Datasets	CDISC SEND	3.0
Controlled Terminology	CDISC SEND Controlled Terminology	2015-06-26
Data Definition file	CDISC DEFINE.XML	1.0

3.2 Rationale for Standards Selection

The Tabulation Datasets Standard or Dictionary used was the latest version available in the Provantis Centrus SEND Solution used for dataset generation.

The Controlled Terminology used was the most current version in the Provantis Centrus SEND solution used at the time of dataset creation; however, translations were manually checked against the most recent published CDISC SEND Controlled Terminology which could not be up versioned on the system at time of dataset preparation.

The Data Definition file is the version currently supported by ppsDefine (part of the Provantis Centrus system which is used to create the define.xml file).



Overview of Content:

3. Standards, Formats, Terminologies & their Versions (cont'd)



- Explain any Terminology used in the study which does not match with [CDISC Controlled Terminology](#), where CDISC CT is required by SEND IG
 - Nonstandard terminology should be [submitted to NCI](#) as a *CDISC New Term Request* for consideration, and explained in nSDRG:

3.3 Nonstandard Terminology

Controlled Terminology for 2015-06-26 was used for this submission. The following extensions to Controlled Terminology [CodeLists](#) were used. The terms will be submitted to NCI for consideration unless already present in the current Controlled Terminology version release:

Dataset Name	Variable	CodeList	Term Used	Meaning
EG	EGTEST	EGTEST	QTcV Interval, Aggregate	A QT aggregate interval corrected for heart rate using the Van der Water's correction formula.
LB	LBTEST	LBTEST	N-acetyl-B-D-glucosaminidase/Creatinine	A relative measurement (ratio or percentage) of the N-acetyl-beta-D-glucosaminidase (enzyme) to creatinine in a biological specimen.
LB	LBTESTCD	LBTESTCD	NAGASECR	A relative measurement (ratio or percentage) of the N-acetyl-beta-D-glucosaminidase (enzyme) to creatinine in a biological specimen.
LB	LBORRESU	LBORRESU	NONE	No Units for pH
MI	MISPEC	SPEC	SMALL INTESTINE, ILEUM/LARGE INTESTINE, COLON/LARGE INTESTINE, CECUM	A tissue sample that contains the ileum, colon, and cecum.
MA	MASPEC	SPEC	SMALL INTESTINE, ILEUM/LARGE INTESTINE, COLON/LARGE INTESTINE, CECUM	A tissue sample that contains the ileum, colon, and cecum.



Overview of Content:

4. Description of Study Datasets



- Provide a Dataset Summary list of domains submitted
- Provide explanations of datasets as needed (see following slides)

4.1 Dataset Summary

Dataset Name	Dataset Label	Supplemental Qualifiers?	Related Records?	Observation Class
TS	Trial Summary			Trial Design
TE	Trial Elements			Trial Design
TA	Trial Arms			Trial Design
TX	Trial Sets			Trial Design
CO	Comments			Special Purpose
DM	Demographics			Special Purpose
SE	Subject Elements			Special Purpose
EX	Exposure			Interventions
DS	Disposition			Events
BW	Body Weight			Findings
BG	Body Weight Gain			Findings
CL	Clinical Observations			Findings
EG	ECG Test Results			Findings
FW	Food and Water Consumption			Findings
LB	Laboratory Test Results			Findings
MA	Macroscopic Findings	X	X	Findings
MI	Microscopic Findings	X	X	Findings
OM	Organ Measurements			Findings
PC	Pharmacokinetic Concentrations			Findings
PP	Pharmacokinetic Parameters			Findings
POOLDEF	Pool Definition			Special Purpose



Overview of Content:

4. Description of Study Datasets (cont'd)



- For Dataset Explanation, *focus on what is important* on a study-by-study basis.
 - Questions that may be helpful for Dataset Explanations:
 - Are the submitted data taken from an ongoing study? If so, what was the cut point(s)?
 - Were any domains planned but not submitted because no data were collected?
 - Are the submitted data a subset of collected data? (The answer will be “yes” if data were collected but not included in the submission.)
 - Describe Supplemental Qualifier domains, if included in the dataset



Overview of Content:

4. Description of Study Datasets (cont'd)



Example

4.2 Dataset Explanation

4.2.1 EX-Exposure

- The EX domain contains details of the actual dose administered at each dose event.

4.2.2 BW-Body Weight Domain:

- Body weight baseline values are identified as the last weight recorded immediately prior to treatment (i.e. Day 1).

4.2.3 CL-Clinical Observations Domain

- Due to limitations with the Provantis LIMS some clinical observations include additional information e.g. 'Post Dose,' which may be populated in the CLLOC field as an extension of the clinical observation recorded.

4.2.4 LB – Laboratory Test Results Domain

- Baseline values are identified as the last measurement recorded prior to the start of treatment.

4.2.5 EG – ECG Test Results Domain

- Baseline values (EGBLFL) are identified as the last measurement recorded prior to the start of treatment.
- The EG before dosing timepoint is presented as 0H
- The EG post dose 2 to 4 hours after dosing is presented as 2-4 for EGTPT and PT2H for EGELTM

4.3 Use of Supplemental Qualifiers

Dataset Name	Associated Dataset	Qualifiers Used
SUPPMA	Macroscopic Observations	Modifiers that were part of MAORRES which SEND variables have not yet been developed.
SUPPMI	Microscopic Observations	Modifiers that were part of MAORRES which SEND variables have not yet been developed.



Overview of Content:

5. Data Standards Validation Rules, Versions, & Conformance Issues



- Summarize and explain pre-submission validation outcomes
- Recommended to focus on FDA rule conformance even if another validator tool was used (we used pinnacle21 Community)
- All issues not fixed should be explained (Errors **and** Warnings)

Rule	Message	Domain(s)	Count	Explanation
FDAN055	Permissible variable with missing value for all records	SUPPMI	1	Warning accepted – these values are not required.
FDAN055	Permissible variable with missing value for all records	TA	2	Warning accepted – these values are not required.
FDAN055	Permissible variable with missing value for all records	TE	1	Warning accepted – these values are not required.
FDAN212	Duplicate records	FW	6	Warning accepted – duplicate times acceptable
FDAN212	Duplicate records	MA	1	Warning accepted – occurs where a different finding is entered for the same tissue (to be investigated).
FDAN212	Duplicate records	MI	19	Warning accepted – occurs where a different finding is entered for the same tissue (to be investigated).
FDAN212	Duplicate records	PP	54	Warning accepted – same specimen type correctly entered
FDAN216	Unexpected character value in COVAL variable	CO	1	Warning accepted – considered due to a 'newline' in the comment.



Overview of Content:

6. Description of Sponsor Decisions Related to Data Standard Implementations



6.1 Sponsor-Defined Standardization Descriptions

- Any discussions between sponsor and review division on how to represent certain study data in instances where SENDIG or SDTMIG may not provide specific instructions
- Explanation for why certain data elements could not be fully standardized
- Explanations of any custom domains ([PhUSE Wiki - guidance on CD](#))
- Comments on inclusion of any derived values
- Include description of any data that may also be located in another study, such as a shared control group

[There were no custom domains or custom endpoints for the example study shown]



Overview of Content:

6. Description of Sponsor Decisions Related to Data Standard Implementations (cont'd)



6.2 Differences Between SEND Datasets and Study Report

- Data included in report but not in datasets or vice versa
- Explanation for multiple study numbers

6.2 Differences between SEND Datasets and Study Report

- Data in the SEND datasets are an accurate representation of data in the study report for Study No. D30476G, with the following differences noted:
- Ophthalmoscopy observations are presented in the Clinical Observations (CL) domain.
- Electrocardiography results for QRS, PR and RR values are included in the Electrocardiography (EG) domain, however these values are not presented in the study report.
- Electrocardiography results for USUBJID D30476G-1 were repeated on study Day -5 and included in the study Day -6 results for mean data purposes in the study report; however they are presented as Day -5 in the Electrocardiography (EG) domain.
- Electrocardiography results for USUBJID D30476G-19 were repeated on study Day 25 and included in the study Day 24 results for mean data purposes in the study report; however they are presented as Day 25 in the Electrocardiography (EG) domain.
- Electrocardiography results for USUBJID D30476G-22 were repeated on study Day 25 and included in the study Day 24 results for mean data purposes in the study report; however they are presented as Day 25 in the Electrocardiography (EG) domain.
- Electrocardiography results for study Day -1 are flagged as baseline results in the Electrocardiography (EG) domain; however, the pre treatment data recorded during different occasions were averaged in the study report to create baseline values for data comparison.
- Urinalysis results collected on study Days -5, -4, 24 and 25 are present in the Laboratory (LB) domain; however these values are not present in the study report. These have been flagged as excluded results.
- Plasma Concentration results with a value of NR are presented in the Plasma Concentration (PC) domain; however they are presented as NQ in the study report. This is an error in the final report.



Overview of Content:

6. Description of Sponsor Decisions Related to Data Standard Implementations (cont'd)



6.3 Nonstandard Electronic Data Submitted

- Data collected using different terminologies
- Electronic data that do not conform to SEND format
- [nSDRG example - "There were no nonstandard electronic data that were part of this submission."
- Possible examples:
 - Presenting 4-point severity scale in MI domain as unable to map to 5-point as in CT
 - Inclusion of tumor.xpt file



Overview of Content:

6. Description of Sponsor Decisions Related to Data Standard Implementations (cont'd)



6.4 Legacy Data Conversion

- Still unclear exactly what is required
- Studies initiated <17 Dec 2016?
- More relevant for clinical datasets
- TCG "Although not strictly a legacy conversion, for nonclinical studies where data is converted to SEND from a previously established collection system, instances may arise where it is not possible to represent a collected data element as a standardized data element. In these cases, there should be an explanation in the nSDRG as to why certain data elements could not be fully standardized or were otherwise not included in the standardized data submission. As the Study Report should contain a complete representation of the study data, no non-standardized data should be submitted. "
- Our FFU example stated 'Not-Applicable' (I have seen others with a similar statement)
- Pilot 2: Terminology used during data collection was used in the study report with that terminology converted to SEND Controlled Terminology during dataset creation using the Provantis Centrus SEND solution. The codelists are included in the DEFINE.XML file.
- *Possible examples*
 - IF in the future *ever* needed to convert existing SEND v3.0 compliant domain datasets to a higher version of SEND
 - MI severity scores – commenting on whether 'converted' or not



Fit-For-Use (FFU) FDA Review



- GSK participated in the SEND Fit-For-Use (FFU) pilot sponsored by FDA (Sep16)
 - 13 different companies submitted a SEND dataset which was reviewed by FDA and the FDA comments/feedback have now been publically shared (13 Sponsors/CROs with SEND studies)
 - [https://wiki.cdisc.org/display/NSFFUW/Nonclinical+\(SEND\)+Fit+for+Use+Workstream+Home](https://wiki.cdisc.org/display/NSFFUW/Nonclinical+(SEND)+Fit+for+Use+Workstream+Home)
- nSDRG was assessed by FDA as part of this review
- Specific comments for each nSDRG reviewed not available (except for one!)
- Summary nSDRG review comments are available (from the link)



Fit-For-Use (FFU) FDA Review: Specific FDA Reviewer Feedback



Specific Feedback on one nSDRG:

- Reviewer #1: There were no major issues identified with the SDRG
- Reviewer #6: Explanations of warnings were understandable.
- Reviewer #6: The SDRG contained all required sections; the information was sufficient to interpret the SEND dataset.
- Reviewer #3: Agree
- Reviewer #6: Fourteen differences between the study report and dataset were noted
- Reviewer #3: It is useful for the SDRG to point out these differences because the differences are apparent in the data visualizations.
- Reviewer #6: Reporting of the errors in the study report under Section 6.2 is not necessary.



Fit-For-Use (FFU) FDA Review



- The following slides summarising FDA review comments on the (12) nSDRG are taken from the CDISC-PhUSE Fit for Use Pilot presentation which is publically available on the Nonclinical (SEND) Fit for Use Workstream wiki....
- [https://wiki.cdisc.org/display/NSFFUW/Nonclinical+\(SEND\)+Fit+for+Use+Workstream+Home](https://wiki.cdisc.org/display/NSFFUW/Nonclinical+(SEND)+Fit+for+Use+Workstream+Home)



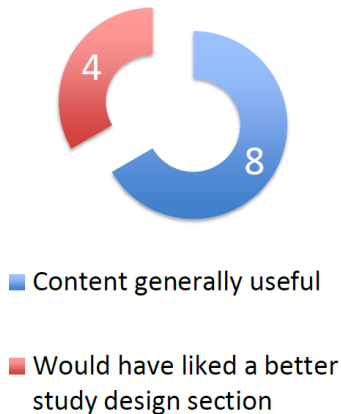
Fit-For-Use (FFU) FDA Review: General FDA Reviewer Feedback



Nonclinical SDRG

FDA

Was the NSDRG useful?



- In general, NSDRG content was very useful
- Most of the comments focused on the clarity and content of the study design section – it is very useful for quickly becoming familiar with a study

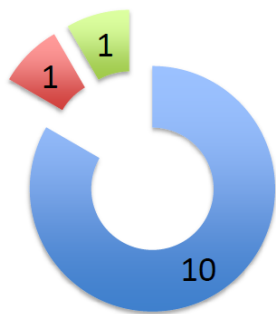


Fit-For-Use (FFU) FDA Review: General FDA Reviewer Feedback (cont'd)



- Level of Information:

Was there too much information?



- Not too much information
- Long list of unused CT extensions
- Unnecessary Study Report errors list

Was there too little information?



- Enough information
- Needed explanation of inconsistencies between the SEND data and study report
- No explanation for missing information or location of data
- Missing trial group explanation



Fit-For-Use (FFU) FDA Review: General FDA Reviewer Feedback (cont'd)



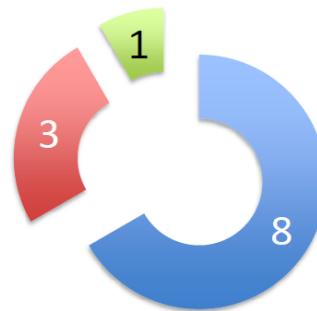
- Specific nSDRG sections:

How was the Study Report vs. Dataset Section?



- Good Study Report vs. Dataset Section
- Failed to mention missing SEND data

How was the Validation Outcomes Section?



- Validation explanations clear
- Difficult to understand
- Validation outcomes not explained



Summary



- FDA reviewers rely heavily on the nSDRG
 - **Study design section** (for fast familiarisation with the study)
 - Study-specific issues, approaches
 - Custom Domains
 - **Study Report vs Dataset Section** - crucial
 - Validation explanation is required!!
- The PhUSE nSDRG link offers a [Template with Guidance](http://www.phusewiki.org/wiki/index.php?title=Nonclinical_Study_Data_Reviewers_Guide)
[http://www.phusewiki.org/wiki/index.php?title=Nonclinical_Study_Data_Reviewers_Guide]
 - help the FDA Reviewer to understand your SEND datasets!
- Please feedback any queries on the Template via the PhUSE Wiki contacts



Current nSDRG PhUSE Working Group Members



Individual	Company	Individual	Company	Individual	Company
Susan DeHaven *	Sanofi	Mark Carfagna	Lilly	Jennifer Abbott	Abbvie
Debra Oetzman *	Instem	Christy Kubin	MPI Research	Ken Hershman	Novartis
Jennifer Feldmann	Instem	Steve Polley	GSK	Anja Metzger	Bayer
Gitte Frausing	Data Standard Decisions	Maureen Rossi	Roche	Susie Lendal Antvorskov	Data Standards Decisions
William Houser	Bristol-Myers Squibb	Tania Strimple	Covance	Matthew Whittaker	FDA
Linda Hunt	Merck	Catherine Roy	Merck	Full Team email list available here Contact us with questions and comments about nSDRG	
Dan Potenta	PDS Lifesciences	Laura Kaufman	Eagle US		
Chris Ely	Pfizer	Michael Rosentreter	Bayer		



Q&A, Discussion.....



- Has anyone here produced a SEND dataset with a nSDRG?
- Did you use the template and guidance from the [PhUSE nSDRG](#) group?
- If so, did this meet your needs? Or is there anything lacking, unclear or can be improved?
- ...or do you have your own specific company template?
- Have you used the Test Submission facility with FDA?
- What were your experiences?



Additional Slides

What Governs SEND?

- [Guidance for Industry - Providing Regulatory Submissions in Electronic Format - Standardized Study Data](#)
 - governs **when** submissions must be in SEND format
- [Study Data Technical Conformance Guide \(March 2017\)](#)
 - supplements the above guidance document.
 - provides specifications, recommendations, and general considerations on how to submit standardized study data using FDA-supported data standards located in the [FDA Data Standards Catalog](#)
- [SEND Implementation Guide](#)
 - v3.0 currently in use
 - v3.1 is available but not being used at this time (not expected to be enforced this year)
 - SEND DART (Development & Reproductive Toxicology) is available for provisional use (not finalized)
- [SEND CDISC Controlled Terminology](#)
 - SEND uses controlled terms for preclinical studies
 - Managed by the National Cancer Institute
 - These are revised every 3 months and made available via the website
 - Note that not all terms are currently controlled
 - Non-neoplastic findings
 - Clinical Observations

nSDRG Overview

2. Study Design



- Further example of a more complex study design table

Study Group SPGRPCD	Trial Arms		Element in each Epoch		Trial Set	
	ARMCD	ARM	Acclimatization	Treatment	SETCD	SET
1	VEH_CTRL	Vehicle Control	Acclimatization	Vehicle Control; 0 mg/kg/dose; ORAL GAVAGE; ONCE DAILY	1.1	Vehicle Control; Non TK; Males
					1.2	Vehicle Control; Non TK; Females
					1.3	Vehicle Control; TK
2	LOW	Low Dose	Acclimatization	GSKXXXXXX; 10 mg/kg/dose; ORAL GAVAGE; ONCE DAILY	2.1	Low Dose; Non TK; Males
					2.2	Low Dose; Non TK; Females
					2.3	Low Dose; TK
3	MID	Mid Dose	Acclimatization	GSKXXXXXX; 50 mg/kg/dose; ORAL GAVAGE; ONCE DAILY	3.1	Mid Dose; Non TK; Males
					3.2	Mid Dose; Non TK; Females
					3.3	Mid Dose; TK
4	HIGH	High Dose	Acclimatization	GSKXXXXXX; 100 mg/kg/dose; ORAL GAVAGE; ONCE DAILY	4.1	High Dose; Non TK; Males
					4.2	High Dose; Non TK; Females
					4.3	High Dose; TK

4. Description of Study Datasets (continued)



- For Dataset Explanation, *focus on what is important* on a study-by-study basis.
 - Previous agreements with regulatory agencies on specific representation of study data not specified in current implementation guides (e.g. SEND IG v. 3.0).
 - Organization of content (e.g. custom endpoints) for which the content is very specific to the study.
 - Inclusion of important non-standard variables supporting key analysis in an associated supplemental qualifier dataset.
 - Description of notable, sponsor defined uses of category and sub-category.
 - Descriptions of derivations/deviations/amendments that may benefit from additional detail beyond that included in define.xml.
 - Description of criteria used to split datasets and the content of the split datasets. (Splitting datasets is expected to be very rare.)
 - Descriptions of relationships to other domains that are documented in RELREC.



Additional examples of Dataset explanations (from independent pilot to be sent to FDA)

4.2.1 EX-Exposure

- The EX domain contains details of the actual volume and the calculated dose administered at each dose event.
- The EXROUTE variable includes method detail (i.e., ORAL GAVAGE) due to limitations with Provantis LIMS.

4.2.2 BW - Body Weight Domain:

- Body weight baseline values are identified as the last weight recorded immediately prior to treatment (i.e., Day 1).

4.2.3 CL - Clinical Observations Domain

- Due to limitations with the Provantis LIMS, some clinical observations include additional information as an extension of the clinical observations recorded, which are populated in the CLLOC field as follows: Yellow foamy, With food, With dose-like material, White foamy, Post dose, Orange liquid, In pen/cage, Fruit remaining and >Half portion remaining. Similarly, the CLSEV are populated as follows: 1 area, 2 areas, 3 areas, >3 areas and At Dosing.
- Food consumption is recorded in the CL domain as a qualitative assessment and is limited to recording 'Eating reduced' (if number of biscuits remaining is greater than or equal to half) and/or 'Fruit remaining' (if any) from the daily allotment.

4.2.4 LB – Laboratory Test Results Domain

- Baseline values (LBBLFL) are identified as the last measurement recorded prior to the start of treatment (i.e., Day -8).

4.2.5 EG – ECG Test Results Domain

- Baseline values (EGBLFL) are identified as the last measurement recorded for each time point (EGELTM) prior to the start of treatment (i.e., Day -13).
- ECG waveforms not obtained due to the loss of the telemetry signal are indicated as NOT DONE and SIGNAL LOST for variables EGSTAT and EGREASND, respectively.
- Qualitative results (INTP) collected over a period of time (e.g., 1-2 hours) are presented for the end of the time period (e.g., 2 hours).

4.2.6 PC – Pharmacokinetics Concentrations Test Results Domain

- Baseline values (PCBLFL) are identified as the 0 hour timepoint for Day 7 and PCORRES for this data is derived. The 0 hour timepoint on Day 7 uses the concentration value from the 24 hour timepoint on Day 7 for toxicokinetic analysis as a predose sample was not required by protocol.
- PCRFTDTC variable is not populated because dosing information is collected on a different LIMS from Pharmacokinetics Concentrations and our systems do not currently support population of this variable across systems.