FDA Study Data Standardization Plan and Study Data Rewiewer's Guide

Daniel Lee Potenta, August, 2016

Study Data Standardization Plan

The Standardization Plan should include, but is not limited to the following:

1. List of the planned studies
2. Type of studies (e.g., phase I, II or III)
3. Study designs (e.g., parallel, cross-over, open-label extension)
4. Planned data standards, formats, and terminologies and their versions
5. List of and justification for studies that may not conform to the standards

Study Data Standardization Plan

Purpose:

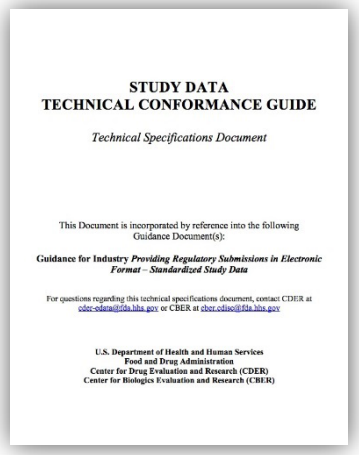
1. To serve as a vehicle to engage FDA in discussions on data standards
2. To inform FDA of what they can expect in terms of data standardization
3. To drive decisions and discussions on legacy data

SDSP applies to an entire program, both clinical and nonclinical. It is best initiated at the preIND stage, and updated throughout the program lifecycle.

SDSP lists individual clinical and nonclinical studies, applicable data standards, and potential conformance issues.

SDSP template, Sponsor Implementation Guide, Completion Guide, and examples can be found at:

http://phusewiki.org/wiki/index.php?title=Study_Data_Standardization_Plan_(SDSP)



STUDY DATA
TECHNICAL CONFORMANCE GUIDE

Technical Specifications Document

This Document is incorporated by reference into the following Guidance Document(s):

Guidance for Industry *Providing Regulatory Submissions in Electronic Format – Standardized Study Data*

For questions regarding this technical specifications document, contact CDER at cder-edata@fda.hhs.gov or CBER at cber.cdisc@fda.hhs.gov

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

SDSP Example

Study Data Standardization Plan

<Sponsor Name>
<Name of Product>
<Indication>
<IND>

Version yyyy-mm-dd

SDSP (cont.)

Study Data Standardization Plan

Contents

1. Introduction 3
1.1 Purpose 3
1.2 Scope 3
1.3 Acronyms 3
1.4 Definitions 4
2. General Sponsor Information 5
3. Product Information 5
4. List of Studies and Standards 6
4.1 Nonclinical 6
4.2 Clinical 7
5. Non-Conformance to Supported Standards Justification 9
6. FDA Data Standards Discussions 10
7. References 11

SDSP (cont.)

<Name of Product> <Indication> Study Data Standardization Plan

1. Introduction

1.1 Purpose

The purpose of the Study Data Standardization Plan (SDSP) is to establish and document a plan for describing the data standardization approach for studies within a development program. The Study Data Standardization Plan (Standardization Plan) assists FDA in identifying potential data standardization issues early in the development program[1].

1.2 Scope

The scope of this document is for use with FDA submissions only. After it has been approved for use, the scope could expand to be used with other regulatory agencies after the proper discussions.

The SDSP is intended to include historical, current, and planned information about the development of the compound and indication. Multiple plans are permissible within a compound. It will be updated and maintained throughout the development of the compound, as new studies are planned or as the data standardization strategy evolves.

Standards that are currently available in the Data Standards Catalog[2] are the basis for which standards are listed.

1.3 Acronyms

Acronym	Translation
ADaM	Analysis Data Model
ISE	Integrated Summary of Efficacy
ISS	Integrated Summary of Safety
LOINC	Logical Observation Identifiers Names and Codes
NDF-RT	National Drug File – Reference Terminology
SDSP	Study Data Standardization Plan
SDTM	Study Data Tabulation Model
SEND	Standard for Exchange of Nonclinical Data
SNOMED	Systematized Nomenclature of Medicine
UNII	Unique Ingredient Identifiers
WHO-DD	WHO Drug Dictionary

SDSP (cont.)

<Name of Product> <Indication> Study Data Standardization Plan

1.4 Definitions

	Term	Definition
Nonclinical	Planned Study	Study Start Date has not been achieved.
	Ongoing Study	Study Start Date has been achieved and the study is not completed.
	Completed Study	The final report has been signed by the study director. (Comparison Chart of FDA, Environmental Protection Agency (EPA), Organization for Economic Co-operation and Development (OECD)[3])
	Study Start Date	The date on which the protocol is signed by the study director. Also known as Study Initiation Date. (Comparison Chart of FDA, Environmental Protection Agency (EPA), Organization for Economic Co-operation and Development (OECD)[3]) (CDISC SEND Controlled Terminology, STSTDTC[5])
	No Electronic Data	Study data is not available electronically.
Clinical	Planned Study	Study Start Date has not been achieved.
	Ongoing Study	1 or more patients is enrolled in the clinical trial[4] and the study is not completed.
	Completed Study	The final subject was examined or received an intervention for the purposes of final collection of data for the primary outcome, whether the clinical trial concluded according to the prespecified protocol or was terminated.[4]
	Study Start Date	The earliest date of informed consent among any subject (Date/Time of Informed Consent, RFICDTC) that enrolled in the study. For studies conducted without informed consent (ie. emergency use) use the date of treatment. Dates for subjects who were screen failures are not included. (CDISC SDTM Controlled Terminology, SSTDTC[6])
Nonclinical & Clinical	Legacy Data	Study data that does not conform to the standards by the date of requirement specified in the published Data Standards Catalog[2].

SDSP (cont.)

2. General Sponsor Information

- Name of Product
- Indication
- IND
- Sponsor Name
- Sponsor Contact
- Sponsor Contact Email

3. General Product Information

SDSP (cont.)

4. List of Studies and Standards

4.1 Nonclinical

Study Identifier	Brief Title	Study Type	Study Status	Study Start Date	Exchange Standards	Terminology Standards
		REPEAT DOSE TOXICITY PRIMARY PHARMACODYNAMICS PHARMACODYNAMIC DRUG INTERACTIONS PHARMACOKINETIC DRUG INTERACTIONS (Please see Completion Guidelines for more information)	COMPLETED ONGOING PLANNED	YYYY-MM-DD <forecasted FPV>	LEGACY SEND IG <version> tumor.xpt define.xml <version> No Electronic Data	LEGACY CDISC SEND Terminology <date> NONE

SDSP (cont)

4.2 Clinical

Study Identifier	Brief Title	Study Design	Study Status	Study Start Date	Exchange Standards	Terminology Standards
<Phase <x>> <Interventional/Observational/Expanded Access> Studies - <indication>						
		Interventional: *Allocation,* *Control Group,* *Intervention Model,* *Masking,* *Study Classification,* *Primary Purpose* **Observational:** *Time Perspective,* *Biospecimen Retention,* *Target Follow-Up,* *Duration* **Expanded Access:** *None – This column is blank* Refer to clinicaltrials.gov study design permissible values[7] See Completion Guidelines for more information	COMPLETED ONGOING PLANNED	YYYY-MM-DD <forecasted FPV>	ANALYSIS LEGACY ADaM <version>/ ADaM IG <version> ADaM define.xml <version> TABULATIONS LEGACY SDTM <version>/ SDTM IG <version> SDTM define.xml <version>	CDISC SDTM Terminology <date> MedDRA (Adverse Events) <version> WHO-DD (Medications) <version> LOINC (Lab Test Terms) <version> SNOMED CT (Indication) <version> NDF-RT (Pharm Class) <version>

5. **Non-Conformance to Supported Standards Justification**
6. **FDA Data Standards Discussions**

SDSP (cont.)
7. References

- Study Data Technical Conformance Guide
  [http://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm]
- Data Standards Catalog
  [http://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm]

  Comparison Chart of FDA, Environmental Protection Agency (EPA), Organization for Economic Co-operation and Development (OECD)
  [http://www.fda.gov/ICECI/EnforcementActions/BioresearchMonitoring]
- US Public Law 110-85, Title VIII, Section 801
  [http://www.gpo.gov/fdsys/pkg/PLAW-110publ85/pdf/PLAW-110publ85.pdf]
- CDISC SEND Controlled Terminology
  [http://evs.nci.nih.gov/ftp1/CDISC/SEND]
- CDISC SDTM Controlled Terminology
  [http://evs.nci.nih.gov/ftp1/CDISC/SDTM]
- Clintrials.gov Study Design Terminology
  [https://prsinfo.clinicaltrials.gov/definitions.html#StudyDesign]

SDRG – Study Data Reviewer Guide

SOME SDRG FAQs

- What is the purpose of an SDRG?
  - Aid FDA reviewer by providing context for SEND datasets beyond what is already provided in define file
  - Highlight any important differences between report & SEND submission
- Is the inclusion of an SDRG with each SEND submission a *binding* requirement?
  - No. It is highly *recommended* by FDA in their Technical Conformance Guidance, which contains nonbinding recommendations.
- Will the FDA / PhUSE SDRG template become a *required* format?
  - Highly unlikely
  - The template will be a *recommended* format

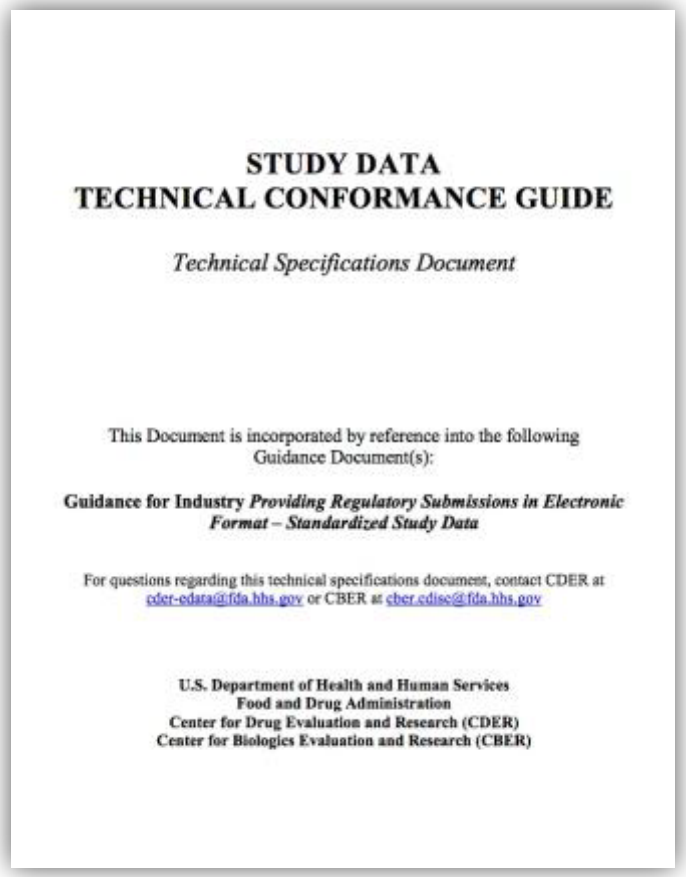
SOME More SDRG FAQs

- Does the FDA expect SEND submissions to include SDRGs while the SDRG format is under review?
  - Yes
- Does the FDA expect Trial SEND submissions to include SDRGs?
  - "That would be helpful. They will get read."

  August 20, 2015 face-to-face meeting between CDISC Core Team and FDA eDATA representative, Lisa Lin
- How does the SDRG differ from the Study Data Standardization Plan (SDSP)?
  - SDSP is submitted at IND stage and covers standardization plans for an entire submission. It is updated as the program matures.
  - SDRG is submitted with each standardized study.

SDRG Table of Contents from PhUSE Template

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
6. Sponsor Decisions Related to Data Standards Implementations



STUDY DATA
TECHNICAL CONFORMANCE GUIDE

Technical Specifications Document

This Document is incorporated by reference into the following Guidance Document(s):

Guidance for Industry *Providing Regulatory Submissions in Electronic Format – Standardized Study Data*

For questions regarding this technical specifications document, contact CDER at cder-edata@fda.hhs.gov or CBER at cber.cdisc@fda.hhs.gov

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

1. Introduction

- Study ID Information
- SEND dataset creation process
- Statement that SEND datasets accurately represent data in the study report and, if needed, where in SDRG any differences are noted

1. Introduction

This document provides context for the SEND tabulation datasets and terminology for Study 54321, in addition to what is provided in the define.xml file, to facilitate the FDA reviewer's and data manager's use of the datasets. It also includes a summary of SEND dataset conformance findings.

1.1 Study Protocol Title, Number, and Report Version

Study Title	A Thirteen-Week Oral Toxicology Study in Dogs with C1234 followed by a Two-Month Recovery Period
Study Number	54321
Report Version	Final. There have been no report amendments.

1.2 Summary of SEND Dataset Creation Process

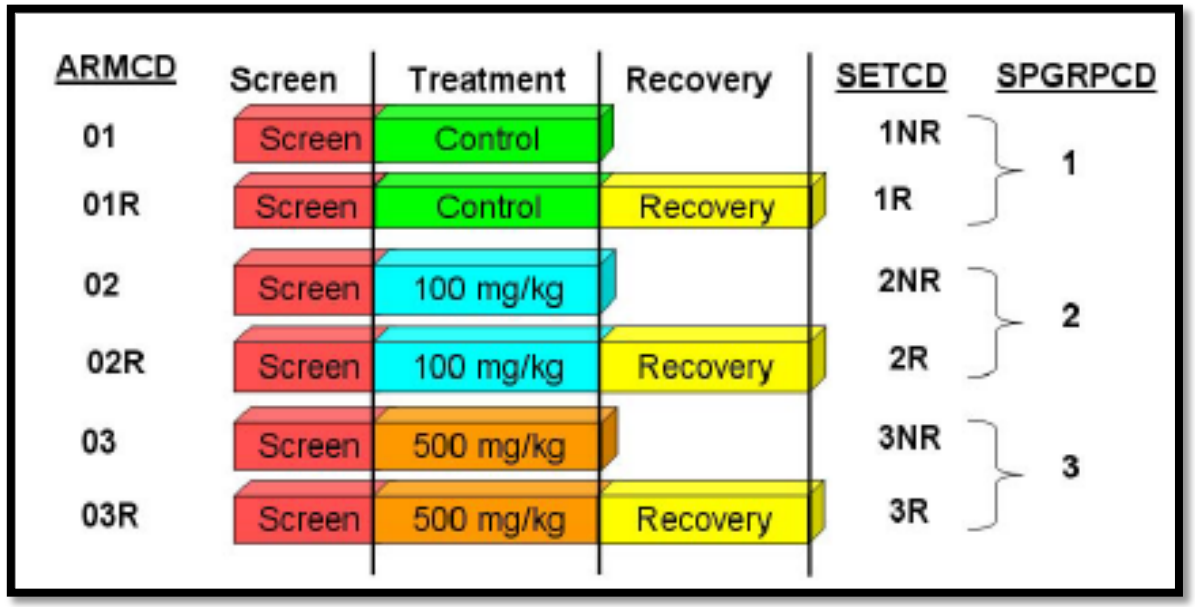
All in-life, clinical pathology, and postmortem data were collected using LIMS 1 (Vendor). Bioanalytical data (plasma drug concentrations) were determined using LIMS 2 (Vendor). Toxicokinetic parameters were calculated using LIMS 3 (Vendor). Input from the each of the LIMS via LIMS-specific adaptors was processed by SEND solution XXX (Vendor) to produce one integrated SEND dataset, define.xml and PDF files, and a validation report. SEND solution XXX and the LIMS-specific adaptors are Part 11 compliant.

1.3 SEND Dataset Verification

Data in the SEND datasets are an accurate representation of data in the study report for Study No. 54321. Any differences between the datasets and the report are described in section 6.2.

2. Study Design

- Provide test or visual representation of SEND Trial Design to help FDA reviewer bridge study design description in the report with SEND Trial Design datasets



3. Standards, Formats, Terminologies and Their Versions

- Table of standards used in SEND data package
- Justify standards and versions selection
- Explain any terminology that does not conform to CDISC CT where CDISC CT is required by SENDIG

3.1 Standards Used

Dataset Component	Standard or Dictionary	Version
Tabulation Datasets	CDISC SEND	3.0
Data Definition File	CDISC DEFINE.XML	1.0
Controlled Terminology (CT)	CDISC SEND CT	2016-03-25

3.2 Rationale for Standards Selection

The standards and versions selected were the most current ones listed in FDA's Study Data Standards Catalog at the time of dataset creation.

3.3 Nonstandard Terminology

Nonstandard terminology was used in the EG domain as shown in the table following.

Dataset Abbreviation	Variable	Term Used	Meaning
EG	EGTEST	Beat-to-beat QT/TQ ratio	A measure of the ability of the heart to recover from one beat to the next by examining the relationship between action potential duration (QT interval) and diastolic interval (TQ) through the ECG.

4. Description of Study Datasets

- Provide a Dataset Summary of all SEND domains submitted

4.1 Dataset Summary

Dataset	Dataset Label	Supplemental Qualifiers?	Related Using RELREC?	Observation Class
TA	Trial Arms			Trial Design
TE	Trial Elements			Trial Design
TS	Trial Summary			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
FW	Food & Water Cons			Events
LB	Laboratory Test Results			Findings
MA	Macroscopic Findings	X	X	Findings
MI	Microscopic Findings	X	X	Findings
OM	Organ Measurements			Findings
PC	Pharmacokinetic Conc			Findings
PP	Parmacokinetic Parameters			Findings
POOLDEF	Pool Definitions			Findings

4. Description of Study Datasets (continued)

- Provide explanations as needed, such as:
  - agreements with review division
  - custom domains
  - nonstandard variables that may support key study analyses
- Focus on ***what is important on a study-by study basis***

4.2 Dataset Explanations

4.2.1 DS Domain

The DSDECOD of UNPLANNED TERMINAL SACRIFICE was used for animals in the high dose treatment group that was terminated early by protocol amendment. Other animals in that group were terminated prior to issuance of the protocol amendment and were assigned a DSDECOD of MORIBUND SACRIFICE.

4.2.2 LB Domain

Urinalysis parameters were characterized in the raw data electronic extracts as either *urinalysis* for volume, protein, or osmolality or *urinalysis excretion* for sodium, potassium, and chloride. The *terms urinalysis* and *urinalysis excretion* were mapped to LBCAT.

4.2.2 MI Domain

RELREC was not used to relate findings in the MI and MA domains because there were no correlative microscopic and macroscopic findings.

4. Description of Study Datasets (continued)

- Questions / items to consider for dataset explanations:
  - Are the submitted data from an ongoing study? If so, what was the cut point?
  - Are the submitted data a subset of collected data? (An example would be a repeat-dose toxicology study with an *in vivo* genotoxicity component)
  - Were the SEND datasets used for the analysis and interpretation in the report?
  - Were any domains planned but not submitted because no data were collected?
  - Nonstandard variables that may support key study analyses

- Describe Supplemental Qualifier Domains, if included

Dataset Name	Associated Dataset	Qualifiers Used
SUPPMA	MA Macroscopic Findings	Modifiers that were part of MAORRES for which SEND variables have not yet been developed
SUPPMI	MI Microscopic Findings	Modifiers that were part of MIORRES for which SEND variables have not yet been developed

5. Data Standards Validation Rules, Versions, and Conformance

- Summarize and explain validation results for Errors and Warnings
- Focus on conformance to published FDA SEND validation rules

5.1 Validation Method

OpenCDISC Validator Version 2.0.1, which includes all FDA SEND validation rules Version 2.0, was used to evaluate conformance to SEND 3.0.

5.2 Validation Outcome Summary

Of a total of 31,682 records, there were 0 errors and 1807 warnings. None of the warnings were relevant to this SEND submission for the reasons provided in section 5.4.

5.3 Errors

No errors were reported.

5.4 Warnings

The Warnings for Study 54321 resulted from a small number of FDA SEND validation rules as shown in the table following. None of these warnings are relevant to this study for the reasons discussed in the table following.

5. Data Standards Validation Rules, Versions, and Conformance (continued)

FDA Rule 2.0	Message	Domain	Count	Explanation
FDAN212	Duplicate records	FW	1347	The validation rule is incorrectly applied because it does not take FWORRES into consideration.
FDAN154	Missing value for LBORRESU when LBORRES is provided	LB	79	The value for LBORRES is albumin / globulin ratio, which is not associated with units. Accordingly, LBORRESU should not be populated, and the validation rule is incorrectly applied.
FDAN164	Missing value for LBSTRESU when LBSTRESC is provided	LB	79	The value for LBSTRESC is albumin / globulin ratio, which is not associated with units. Accordingly, LBSTRESU should not be populated, and the validation rule is incorrectly applied.
FDAN033	SEND/dataset variable label mismatch	LB	1	The variable in question is LBSTNRC (Reference Range for Character Result-Standard Unit), and it was used in the LB dataset in accord with SENDIG 3.0. Accordingly, the validation rule was incorrectly applied.
FDAN341	PCORRESU and PCSTRESU values not found in Unit extensible codelist	PC	276	The PCORRESU and PCSTRESU values in question are pg/mL. The correct CT codelist for these variables is PKUNIT and not Unit. The validation rule is therefore incorrectly configured and does not apply here.
FDAN212	Duplicate records	PP	25	The validation rule is incorrectly applied because it does not take PPORRES into consideration.

6. Description of Sponsor Decisions Related to Data Standards Implementations

6.1 Sponsor-Defined Standardization Descriptions

- Explanation for why certain data elements could not be fully standardized
- 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

6.2 Differences Between SEND Datasets and Study Report

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

6.3 Nonstandard Electronic Data Submitted

- Data collected using different terminologies
- Electronic data that do not conform to SDTM format

6.2 Differences Between SEND Datasets and Study Report - Example

Situation:	One protocol (Study A), one report (Study A), with toxicology portion of study entered into LIMS as “Study A_TOX” and toxicokinetics portion entered into LIMS as “Study A_TK”
Mapping Challenge:	Two separate raw data extracts from LIMS
Solution:	Submitted 2 separate SEND datasets, 1 for toxicology component, 1 for TK component, used Trial Summary Parameter ASOCSTDY to link both datasets. Summarized this solution in SDRG.
FDA Feedback:	FDA has used this SDRG as an example of a good SDRG. Without the SDRG, the datasets could not have been processed. With the SDRG, the datasets were validated and analyzed the first time.

SENDIG 3.1 was released by CDISC July 8, 2016

- The most significant changes since the prior version, v3.0, include:
- • Realignment with SDTM v1.5.
- • Reclassified the use of VISITDY from Expected to Permissible and added three new variables to relevant domains (--USCHFL, --NOMLBL, --NOMDY)
- • New FOCID variable added to several domains (EX, CL, MA, and MI). FOCID, described further in Section 4.2.7, is available to all general observation classes to specify a study-specific point of interest.
- • Updated Microscopic Findings domain (Section 6.3.8); added three new variables (FOCID, --CHRON, --DISTR).
- • Updated Vital Signs domain (Section 6.3.15). With the addition of new domains for nonclinical safety pharmacology study data, CV and RE, the Vital Signs domain was revised, blood pressure measurements were moved to the Cardiovascular (CV) domain, and respiratory rate was moved to the Respiratory (RE) domain. For nonclinical study data, the Vital Signs domain is now intended to hold vital signs measurements, such as body temperature, that are not otherwise covered in domains for respiratory and cardiovascular test data.
- • Updated ECG Test Results domain (Section 6.3.16); added two new Timing variables (--STINT and --ENINT).
- • New Cardiovascular domain (Section 6.3.17) including blood pressure measurements, which were previously submitted in Vital Signs domain.
- • New Respiratory domain (Section 6.3.18) including respiratory rate, which was previously submitted in Vital Signs domain.

- A detailed list of changes between versions is provided in Appendix D.
- A list of approved SDTM variables that should not be used in SEND datasets is provided in Appendix E.


Questions