Specifications for Preparing and Submitting Summary Level Clinical Site Data for CDER's Inspection Planning

Revision History

Date	Version	Summary of Changes
11/08/11	1.0	Original Version
10/15/12	1.1	Updated example values in Appendix 2
11/7/12	1.2	Updated Country codes from alpha-2 to alpha-3

TABLE OF CONTENTS

I. ORGANIZATION OF THE SITE DATASET 1

II. VARIABLES AND VARIABLE NAMES FOR SITE-SPECIFIC EFFICACY RESULTS 1

III. CREATING THE DATA FILE (TEMPLATE AND STRUCTURE) 3

A. SUBMISSION TEMPLATE FOR BIORESEARCH MONITORING (BIMO) CLINICAL DATA 3
B. SUBMITTING BIMO CLINICAL DATA IN ECTD FORMAT 3

IV. ELECTRONIC TRANSPORT FORMAT 4

V. IDENTIFICATION OF THE DATASET 4

VI. SUBMISSION OF THE DATASET 4

APPENDIX 1: CLINICAL SITE DATA ELEMENTS SUMMARY LISTING 5

APPENDIX 2: EXAMPLE FOR CLINICAL SITE DATA ELEMENTS SUMMARY LISTING 8

Specifications for Preparing and Submitting Summary Level Clinical Site Data for CDER's Inspection Planning

This document provides current FDA specifications for preparing and submitting a summary level clinical site dataset in electronic form for new drug applications (NDAs), biologics licensing applications (BLAs), and NDA or BLA supplemental applications submitted to the Center for Drug Evaluation and Research (CDER).

I. ORGANIZATION OF THE SITE DATASET

Submit a single summary level clinical site dataset that contains data from all major (e.g., pivotal) studies used to support safety and efficacy in the application, including studies with different treatment indications.

For each major (e.g., pivotal) study used to support safety and efficacy, submit data by clinical site and treatment arm for the intent-to-treat (ITT) population.

For clinical investigator sites involved in multiple studies in support of an application, provide the data independently for each study within the dataset.

II. VARIABLES AND VARIABLE NAMES FOR SITE-SPECIFIC EFFICACY RESULTS

For each study and investigator site, use the following variables associated with efficacy and their variable names:

- Treatment Efficacy Endpoint (TRTEFFE) – the summary statistic for each primary efficacy endpoint, by treatment arm (see below for examples of summary statistics according to different types of efficacy endpoints)
- Treatment Efficacy Endpoint Standard Deviation (TRTEFFS) – the standard deviation of the summary statistic (TRTEFFE) for each primary endpoint, by treatment arm
- Site-specific Treatment Effect (SITEEFFE) – the treatment effect should be reported using the same representation as reported for the primary efficacy analysis
- Site-specific Treatment Effect Standard Deviation (SITEEFFS) – the standard deviation of the site-specific treatment effect (SITEEFFE)
- Endpoint (endpoint) – a plain text label that describes the primary endpoint as described in the data definition file data dictionary included with each application.
- Treatment Arm (ARM) – a plain text label for the treatment arm that is used in the Clinical Study Report

In addition, for studies whose primary endpoint is a time-to-event endpoint, include the following data element:

- Censored Observations (CENSOR) – the number of censored observations for the given site and treatment.

If a study does not contain a time-to-event endpoint, record this data element as a missing value.

To accommodate the variety of endpoint types that can be used in analyses, reference the following endpoint type definitions when tabulating the site-specific summary statistic by treatment arm, "TRTEFFE."

- Discrete Endpoints – endpoints based on efficacy observations that can take on a discrete number of values (e.g., binary, categorical). Summarize discrete endpoints by an event frequency (i.e., number of events), proportion of patients with an event, proportion of patients responding to treatment, or similar method at the site for the given treatment.
- Continuous Endpoints – endpoints based on efficacy observations that can take on an infinite number of values. Summarize continuous endpoints by the mean, median, or other distributional quantile of the observations at the site for the given treatment.
- Time-to-Event Endpoints – endpoints where the time to occurrence of an event is the primary efficacy measurement. Summarize time-to-event endpoints by two data elements: the number of events that occurred (TRTEFFE) and the number of censored observations (CENSOR).
- Other – if the primary efficacy endpoint cannot be summarized in terms of the previous guidelines, a single or multiple values with precisely defined variable interpretations should be submitted as part of the dataset.

In all cases, the endpoint description provided in the "endpoint" plain text label should be expressed clearly to interpret the value provided in the (TRTEFFE) variable.

The site-specific treatment effect (SITEEFFE) should be summarized in terms of the primary efficacy analysis (e.g., difference of means, difference of proportions, odds ratio, hazard ratio, etc.) and should be defined identically for all records in the dataset regardless of treatment.

The summary level clinical site dataset should be accompanied by a data definition file. A sample define file for a dataset is presented in Appendix 1.

III. CREATING THE DATA FILE (TEMPLATE AND STRUCTURE)

A. Submission Template for Bioresearch Monitoring (BIMO) Clinical Data

A sample summary level clinical site data submission using the variables identified in Appendix 1 is provided in Appendix 2.

B. Submitting BIMO Clinical Data in the eCTD Format

A summary level clinical site dataset submitted with an application in the eCTD format belongs in Module 5 – Clinical Study Reports.

- Construct a BIMO study tagging file (STF) and place it in Module 5.3.5.4 Other Study reports and related information. The study ID for this STF is "bimo."

- For the site-level dataset, use the filename "clinsite.xpt."

- Link the site-level dataset files into this BIMO STF, using file tags indicated below.

STF File Tag	Used For	Allowable File Formats
data-listing-dataset	Site-level datasets, across studies	.xpt
data-listing-data-definition	Define file	.pdf

- Within the directory structure, the site-level dataset should be placed in the M5 folder as follows:

References for eCTD submissions:

- eCTD Backbone Specification for Study Tagging Files v. 2.6.1 (http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM163560.pdf)

- FDA eCTD Web page (http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm153574.htm)

- For general help with eCTD submissions: ESUB@fda.hhs.gov

IV. ELECTRONIC TRANSPORT FORMAT

The summary level clinical site data should be submitted in SAS transport file format (*.xpt). See "Study Data Specifications." [1]

V. IDENTIFICATION OF THE DATASET

For the leaf representing the data set, please clearly identify it in the leaf title such as "summary level clinical site date for inspection."

VI. SUBMISSION OF THE DATASET

Please see the Specification for Transmitting Electronic Submissions using eCTD Specifications (http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM163567.pdf) for details on electronic transmission or physical media submissions.

[1] Available at http://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM312964.pdf.

Appendix 1: Clinical Site Data Elements Summary Listing

Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Notes or Description	Sample Value
1	STUDY	Study Number	Char	String	Study or trial identification number.	ABC-123
2	STUDYTL	Study Title	Char	String	Title of the study as listed in the clinical study report (limit 200 characters)	Double blind, randomized placebo controlled clinical study on the influence of drug X on indication Y
3	DOMAIN	Domain Abbreviation	Char	String	Two-character identification for the domain most relevant to the observation. The Domain abbreviation is also used as a prefix for the variables to ensure uniqueness when datasets are merged.	DE
4	SPONNO	Sponsor Number	Num	Integer	Total number of sponsors throughout the study. If there was a change in the sponsor while the study was ongoing, enter an integer indicating the total number of sponsors. If there was no change in the sponsor while the study was ongoing, enter “1.”	1
5	SPONNAME	Sponsor Name	Char	String	Full name of the sponsor organization conducting the study at the time of study completion, as defined in 21 CFR 312.3(a).	DrugCo, Inc.
6	IND	IND Number	Num	6 digit identifier	Investigational New Drug (IND) application number. If study not performed under IND, enter -1.	010010
7	UNDERIND	Under IND	Char	String	Value should equal "Y" if study at the site was conducted under an IND and "N" if study was not conducted under an IND (i.e., 21 CFR 312.120 studies).	Y
8	NDA	NDA Number	Num	6 digit identifier	FDA new drug application (NDA) number, if available/applicable. If not applicable, enter -1.	021212
9	BLA	BLA Number	Num	6 digit identifier	FDA identification number for biologics license application, if available/applicable. If not applicable, enter -1.	123456
10	SUPPNUM	Supplement Number	Num	Integer	Serial number for supplemental application, if applicable. If not applicable, enter -1.	4
11	SITEID	Site ID	Char	String	Investigator site identification number assigned by the sponsor.	50
12	ARM	Treatment Arm	Char	String	Plain text label for the treatment arm as referenced in the clinical study report (limit 200 characters).	Active (e.g., 25mg), Comparator drug product name (e.g., Drug x), or Placebo
13	ENROLL	Number of Subjects Enrolled	Num	Integer	Total number of subjects enrolled at a given site by treatment arm.	20
14	SCREEN	Number of Subjects Screened	Num	Integer	Total number of subjects screened at a given site.	100

Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Notes or Description	Sample Value
15	DISCONT	Number of Subject Discontinuations	Num	Integer	Number of subjects discontinuing from the study after being enrolled at a site by treatment arm as defined in the clinical study report.	5
16	ENDPOINT	Endpoint	Char	String	Plain text label used to describe the primary endpoint as described in the Define file included with each application (limit 200 characters).	Average increase in blood pressure
17	ENDPTYPE	Endpoint Type	Char	String	Variable type of the primary endpoint (i.e., continuous, discrete, time to event, or other).	Continuous
18	TRTEFFE	Treatment Efficacy Endpoint	Num	Floating Point	Summary statistic for each primary efficacy endpoint by treatment arm at a given site.	0, 0.25, 1, 100
19	TRTEFFS	Treatment Efficacy Endpoint Standard Deviation	Num	Floating Point	Standard deviation of the summary statistic (TRTEFFE) for each primary efficacy endpoint by treatment arm at a given site.	0.065
20	SITEEFFE	Site-Specific Treatment Effect	Num	Floating Point	Site-specific treatment effect reported using the same representation as reported for the primary efficacy analysis.	0, 0.25, 1, 100
21	SITEEFFS	Site-Specific Treatment Effect Standard Deviation	Num	Floating Point	Standard deviation of the site-specific treatment effect (SITEEFFE).	0.065
22	CENSOR	Censored Observations	Num	Integer	Number of censored observations at a given site by treatment arm. If not applicable, enter -1.	5
23	NSAE	Number of Non-Serious Adverse Events	Num	Integer	Total number of non-serious adverse events at a given site by treatment arm. This value should include multiple events per subject and all event types (i.e., not limited to only those that are deemed related to study drug or treatment emergent events).	10
24	SAE	Number of Serious Adverse Events	Num	Integer	Total number of serious adverse events excluding deaths at a given site by treatment arm. This value should include multiple events per subject.	5
25	DEATH	Number of Deaths	Num	Integer	Total number of deaths at a given site by treatment arm.	1
26	PROTVIOL	Number of Protocol Violations	Num	Integer	Number of protocol violations at a given site by treatment arm as defined in the clinical study report. This value should include multiple violations per subject and all violation type (i.e., not limited to only significant deviations).	20
27	FINLMAX	Maximum Financial Disclosure Amount	Num	Floating Point	Maximum financial disclosure amount ($USD) by any single investigator by site. Under the applicable regulations (21 CFR Parts 54, 312, 314, 320, 330, 601, 807, 812, 814, and 860). If unable to obtain the information required to the corresponding statements, enter -1.	20000.00

Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Notes or Description	Sample Value
28	FINLDISC	Financial Disclosure Amount	Num	Floating Point	Total financial disclosure amount ($USD) by site calculated as the sum of disclosures for the principal investigator and all sub-investigators to include all required parities. Under the applicable regulations (21 CFR Parts 54, 312, 314, 320, 330, 601, 807, 812, 814, and 860). If unable to obtain the information required to the corresponding statements, enter -1.	25000.00
29	LASTNAME	Investigator Last Name	Char	String	Last name of the investigator as it appears on the FDA 1572.	Doe
30	FRSTNAME	Investigator First Name	Char	String	First name of the investigator as it appears on the FDA 1572.	John
31	MINITIAL	Investigator Middle Initial	Char	String	Middle initial of the investigator, if any, as it appears on the FDA 1572.	M
32	PHONE	Investigator Phone Number	Char	String	Phone number of the primary investigator. Include country code for non-US numbers.	44-555-555-5555
33	FAX	Investigator Fax Number	Char	String	Fax number of the primary investigator. Include country code for non-US numbers.	44-555-555-5555
34	EMAIL	Investigator Email Address	Char	String	Email address of the primary investigator.	john.doe@mail.com
35	COUNTRY	Country	Char	ISO 3166-1-alpha-3	3 letter ISO 3166 country code in which the site is located.	USA
36	STATE	State	Char	String	Unabbreviated state or province in which the site is located. If not applicable, enter NA.	Maryland
37	CITY	City	Char	String	Unabbreviated city, county, or village in which the site is located.	Silver Spring
38	POSTAL	Postal Code	Char	String	Postal code in which site is located. If not applicable, enter NA.	20850
39	STREET	Street Address	Char	String	Street address and office number at which the site is located.	1 Main St, Suite 100

The following is a fictional example of a data set for a placebo-controlled trial. Four international sites enrolled a total of 205 subjects who were randomized in a 1:1 ratio to active or placebo. The primary endpoint was the percent of responders. The site-specific efficacy effect size (SITEEFFE) is the difference between the active and the placebo treatment efficacy result. Note that since there were two treatment arms, each site contains two rows in the following example data set and a total of eight rows for the entire data set.

Appendix 2: Example for Clinical Site Data Elements Summary Listing

STUDY	STUDYTL	DOMAIN	SPONNO	SPONNAME	IND	UNDERIND	NDA	BLA	SUPPNUM	SITEID	ARM	ENROLL	SCREEN	DISCONT
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	001	Active	26	61	3
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	001	Placebo	25	61	4
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	002	Active	23	54	2
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	002	Placebo	25	54	4
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	003	Active	27	62	3
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	003	Placebo	26	62	5
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	004	Active	26	60	2
ABC-123	Double blind…	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	004	Placebo	27	60	1

ENDPOINT	ENDTYPE	TRTEFFE	TRTEFFS	SITEEFFE	SITEEFFS	CENSOR	NSAE	SAE	DEATH	PROTVIOL	FINLMAX	FINLDISC	LASTNAME	FRSTNAME
Percent Responders	Binary	0.48	0.0980	0.34	0.1405	-1	0	2	0	1	-1	-1	Doe	John
Percent Responders	Binary	0.14	0.0694	0.34	0.1405	-1	2	2	0	1	-1	-1	Doe	John
Percent Responders	Binary	0.48	0.1042	0.33	0.1427	-1	3	2	1	0	45000.00	45000.00	Washington	George
Percent Responders	Binary	0.14	0.0694	0.33	0.1427	-1	0	2	0	3	20000.00	45000.00	Washington	George
Percent Responders	Binary	0.54	0.0959	0.35	0.1448	-1	2	2	0	1	15000.00	25000.00	Jefferson	Thomas
Percent Responders	Binary	0.19	0.0769	0.35	0.1448	-1	3	6	0	0	22000.00	25000.00	Jefferson	Thomas
Percent Responders	Binary	0.46	0.0977	0.34	0.1275	-1	4	1	0	0	0.00	0.00	Lincoln	Abraham
Percent Responders	Binary	0.12	0.0625	0.34	0.1275	-1	1	2	0	1	0.00	0.00	Lincoln	Abraham

MINITIAL	PHONE	FAX	EMAIL	COUNTRY	STATE	CITY	POSTAL	STREET
M	555-123-4567	555-123-4560	John@mail.com	RU	Moscow	Moscow	103009	Kremlin Road 1
M	555-123-4567	555-123-4560	John@mail.com	RU	Moscow	Moscow	103009	Kremlin Road 1
	020-3456-7891	020-3456-7890	george@mail.com	GB	Westminster	London	SW1A 2	10 Downing St
	020-3456-7891	020-3456-7890	george@mail.com	GB	Westminster	London	SW1A 2	10 Downing St
	01-89-12-34-56	01-89-12-34-51	tom@mail.com	FR	N/A	Paris	75002	1, Rue Road
	01-89-12-34-56	01-89-12-34-51	tom@mail.com	FR	N/A	Paris	75002	1, Rue Road
	555-987-6543	555-987-6540	abe@mail.com	US	Maryland	Rockville	20852	1 Rockville Pk.
	555-987-6543	555-987-6540	abe@mail.com	US	Maryland	Rockville	20852	1 Rockville Pk.