PharmaSUG 2019 - Paper 162

Multiple Studies BIMO Submission Package – A Programmer's Perspective

Ramanjulu Valluru, Harsha H. Dyavappa, Accenture Inc.

ABSTRACT

As support documentation of its Bioresearch Monitoring (BIMO) activities, FDA's Center for Drug Evaluation and Research (CDER) requests that sponsors of new drug applications (NDAs), biologics license applications (BLAs), and NDA or BLA supplemental applications containing clinical data provides the following three items: Clinical Study-Level Information; Subject-Level Data Line Listings by Clinical Site; Summary-Level Clinical Site Dataset.

Recent papers (Singh Kahlon et al [1], Lin et al [2]), give details on how to create the BIMO submission package containing the three items above. This paper emphasizes multiple studies submissions cases when working with multiple studies to align the submission with FDA's requirements in these situations. It expands on how to create a single clinsite dataset, define.xml and BIMO reviewer's guide instead of one version of each of these documents per study. Specific considerations are given to discuss about a single clinsite dataset; define.xml; and BIMO reviewer's guide used to submit this package in eCTD module 5.3.5.4 submission. We will share our experiences while supporting successful FDA applications for several therapeutic areas on multiple studies working on BIMO site level data.

INTRODUCTION

To protect the safety, integrity and welfare of the study subjects, the US FDA conducts inspection and data audits of Investigator sites through the Office of Scientific Investigations (OSI) for the clinical studies that are registered for New Drug Applications (NDAs), Biologics Licensing Applications (BLAs) and their supplemental applications. To facilitate the inspections and data audits, FDA has issued BIMO submission guidelines for study sponsors to submit documents/datasets to CDER in eCTD format. CDER uses the Summary-Level Clinical Site Dataset as a source data for risk-based modeling to evaluate and select Investigational sites for conducting inspection. The sponsors who can provide accurate and reliable details through BIMO submission can avoid delay on further processes that can impact drug approval timelines. The sponsors internally use BIMO information to track the regulatory compliance of their study sites and investigators, vendors and partners; reliability of the clinical data collected; safety and efficacy of the data as an information to analyze the success of their clinical trials and meeting the study objectives.

The BIMO submission package should contain the following three items:

I. Clinical Study-Level Information
II. Subject-Level Data Line Listings by Clinical Site
III. Summary-Level Clinical Site Dataset

Clinsite dataset labelled as "Summary-Level Clinical Site Dataset" is a unique dataset structured and labelled as per FDA guidelines which includes key information on the safety and primary efficacy endpoints of all the submission ready pivotal studies conducted at each site along with the investigator details.

This paper emphasizes multiple studies submission cases from a programmer point-of-view. It explains how to create a single clinsite dataset; define.xml; BIMO listings and Reviewer's guide that are submitted in eCTD under module 5.3.5.4.

SCOPE OF WORK AND PRE-REQUISITE

The Clinical programmer should create:

I. Clinical Study-Level information
II. Subject-Level Data Line Listings by Clinical Site in pdf format.
III. Summary-Level Clinical Site Dataset in SAS® transport file format named "clinsite.xpt".

IV. Metadata or Define document in pdf format.
V. BIMO Reviewer's Guide

At least the following items must be available before starting the process: Study Protocol, Protocol Amendments, Annotated Case Report form (aCRF), Statistical Analysis Plan (SAP), Listing Mocks, SDTM and ADAM datasets, the Site-level Investigator file (Excel file or dataset) and current FDA's Bioresearch Monitoring Technical Conformance Guide.

I. CLINICAL STUDY-LEVEL INFORMATION

Table 1:
A word document containing all sites will be created using the information from Summary-Level Clinical Site Dataset and convert to PDF format.

Table 2:
By applying SAS programming, for each site, we create a summary table to get the total counts of subjects screened, subjects randomized, subjects treated who prematurely discontinued. ADSL will serve as the source dataset. This summary table can be exported as a word/excel file for proper tabulation and formatting and later converted to a PDF file.

II. SUBJECT-LEVEL DATA LINE LISTINGS BY CLINICAL SITE IN PDF FORMAT

There are 10 unique type listings to be generated at Subject-level, by clinical site. However, if the sponsor believes alternative listings or formats are preferable for their submission, proposed alternatives should be discussed with the Office of Scientific Investigations (OSI) in advance of the application submission, for example, before or during pre-NDA or -BLA meetings. The source data for these listings should be the same as individual clinical study reports (eg: SDTM/ADaM). For multiple studies, these listings should be created separately for each site for each study.

Note: In below tables Row1: Variable Label; Row2: SAS Variable Name; Row3: Example

1. Consented Subjects: Listing includes all subjects that consented to enroll in the study.

Listing 01.001
Listing of Subject Eligibility

Study Site Identifier=001

Subject ID Age/Gender/Race	Planned Treatment	Met Inclusion/Exclusion Criteria	Criteria Number Not Met	Criteria Type Not Met
ADSL.SUBJID/ADSL.AGE/ ADSL.SEX/ADSL.RACE	ADSL.ARM	IE.IESTRESC	Numeric part of IE.IETESTCD	IE.IECAT
001-0001 51/F/White	DRUG A	Yes		

2. **Treatment Assignment:** Listing includes all randomized subjects.

Listing 02.001
Listing for Treatment Assignment (Randomization)

Study Site Identifier=001

Subject ID Age/Gender/Race	Date of Randomization	Planned Treatment
ADSL.SUBJID/ ADSL.AGE /ADSL.SEX/ADSL.RACE	ADSL.RANDDT	ADSL.ARM

001-0001/ 51/F/White	2016-02-15	DRUG A

3. Discontinuations: Listing includes all discontinued subjects irrespective of reasons.

Listing 03.001
Listing of Drop-outs and Discontinuation

Study Site Identifier=001

Subject ID Age/Gender/ Race	Planned Treatment	Study Status	Date of Randomization	Date of Discontinuation	Study Day[1]	Reason
ADSL.SUBJID/ ADSL.AGE/ ADSL.SEX/ ADSL.RACE	ADSL.ARM	DS.DSDECOD= ”COMPLETED” then “Completed“ else “Discontinued”	ADSL.RANDDT	DS.DSSTDTC	DS. DSSTDY	DS. DSDECOD
001-0011 69/F/White	DRUG A	Discontinued	17MAR2016	05APR2016	20	Withdrew Consent

4. StudyPopulation: Listing includes all types of study population subjects to be analyzed and the excluded subjects are reported with reason for exclusion.

Listing 04.001

Evaluable Subjects/Non-Evaluable Subjects and Reason Not Evaluable

Study Site Identifier=001

Subject ID Age/Gender/Race	Planned Treatment	Population Name	Evaluable Y/N	Reason Not Evaluable
ADSL.SUBJID ADSL.AGE /ADSL.SEX/ ADSL.RACE	ADSL.ARM	Safety Intent-to-Treat (ITT) Per-Protocol (PP)	ADSL.SAFFL ADSL.ITTFL ADSL.PPROTFL	DS.DSREAS
001-0001 51/F/White	DRUG A	Safety Intent-to-Treat (ITT) Per-Protocol (PP)	Y Y Y	

5. Inclusion and Exclusion Criteria: Listings include all the subjects who met each inclusionand exclusion criterion as defined in the protocol.

Listing 05.001

Listing of Inclusions and Exclusions Not Met at Screening

Study Site Identifier=001

Subject ID Age/Gender/Race	Met All Inclusion/Exclusion Criteria (Y/N)	Criteria	Criteria Number Not Met	Inclusion/Exclusion Criterion
ADSL.SUBJID/ ADSL.AGE/ ADSL.SEX/ ADSL.RACE	IE.IESTRESC	IE.IECAT	Number part of IE.IETESTCD	IE.IETEST

001-0037 58/M/White	N	Inclusion	03	CRITERIA THAT IS NOT MET

6. Adverse Events: The listing includes all adverse events.

Listing 06.001

Listing of AEs, SAEs, Deaths and Dates

Study Site Identifier=001

Subject ID Age/Gender/ Race	Planned Treatment	System Organ Class/ Preferred Term/ Verbatim Text	Outcome[1]/ Onset Date (Day)/ Resolve Date (Day)[2]/ Duration (Days)	Severity/ Serious	Action Taken[3]/ Relation to Study Drug[4]	AE Death
ADSL.SUBJID /ADSL.AGE/ ADSL.SEX/ ADSL.RACE	ADSL. ARM	AE.AEBODSYS/ AE.AEDECOD/ AE.AETERM	AE.AEOUT/ AE.AESTDTC (AESTDY)/ AE.AEENDTC (AEENDY)/ AEDUR	AE. AESEV/ AE. AESER	AE.AEACN/ AE.AEREL	AE. AESDTH
001-0011 79/F/White	DRUG A	Injury, Poisoning And Procedural Complications/ Procedural Headache/ Headache Post treatment	Recovered/Resolved/ 2016-03-17 (1)/ 2016-03-19 (3)/ 3	Moderate/ N	Dose Not Changed/ Related	N

7. Important Protocol Deviations: The listing should include all important protocol deviations as reported in the NDA or BLA, including a description of the violation/deviation.

Listing 07.001

Listing of Protocol Violations and/or Deviations

Study Site Identifier=001

Subject ID Age/Gender/Race	Planned Treatment	Protocol Deviation Description	Category for Protocol Deviation Description
ADSL.SUBJID/ ADSL.AGE /ADSL.SEX/ ADSL.RACE	ADSL.ARM	DV.DVTERM	DV.DVDECOD
012-0013/79/M/White	DRUG B	Protocol Violation	Protocol Violation

8. Efficacy Endpoints: The listing includes primary and key secondary efficacyparameters or events. For derived or calculated endpoints, the raw data points used to generate the derived or calculated endpoint should be provided.

Listing 08.001

Listing of the Primary and Secondary Endpoint Efficacy Parameters

Study Site Identifier=001

Subject ID Age/Gender/ Race	Planned Treatment	ITT Flag	Per-Protocol Flag	Endpoint Type	Endpoint Description	Value
ADSL. SUBJID/ ADSL.AGE/ ADSL.SEX/ ADSL.RACE	ADSL.ARM	ADSL. ITTFL	ADSL. PPROTFL	Primary Secondary	Primary end point text Secondary end point text	XX XX
001-0001 56/F/White	DRUG A	Y	Y	Primary	Responder analysis of proportion of Randomized subjects with all Target Lesions to be Clear on PLA (PLA=Clear) at Visit 8 (Day 106)	No
				Secondary	Responder analysis of proportion of ITT population subjects with 3 of 4 Target Lesions to be Clear on PLA (PLA=Clear) at Visit 8 (Day 106)	No

9. Concomitant Medications: Listing contain all concomitant medications as required to be collected by the protocol.

Listing 09.001

Listing of Concomitant Medications

Study Site Identifier=001

Subject ID Age/Gender /Race	Planned Treatment	Start Date/ Stop Date	ATC Class Level 1	ATC Class Level 3	Indication	Dose/ Units/ Frequency/ Route	Ongoing
ADSL. SUBJID/ ADSL.AGE/ ADSL.SEX/ ADSL. RACE	ADSL.ARM	ADCM. CMSTDTC/ ADCM. CMENDTC	ADCM.ATC1	ADCM.ATC3	ADCM.CMI NDC	ADCM. CMDOSE/ CMDOSEU CMFREQ/ CMROUTE	ADCM. CMENRTPT
001-0002 47/M/White	DRUG B	1998/	Alimentary Tract and Metabolism	Vitamin A And D, Incl. Combinations of the two	Bone Health	5000/ IU/ Qd/ Oral	Yes

10. Safety Monitoring: Listing includes tests results (e.g., laboratory, electrocardiogram) performed for safety monitoring as defined in the protocol.

Listing 10.001

Listing of Laboratory Test Performed for Safety Monitoring

Study Site Identifier=001

Subject ID Age/Gender/ Race	Planned Treatment	Lab Test	Category for Lab Test	Analysis Visit	Analysis Date	Result/ Standard Units	Normal Range	Lower Limit/ Upper Limit
ADSL.SUBJID ADSL.AGE/ ADSL.SEX/ ADSL.RACE	ADSL. ARM	ADLB. LBTEST	ADLB.LBCAT	ADLB.AVISIT	ADLB.ADT	ADLB. LBSTRESC ADLB. LBSTRESU	ADLB. LBSTN RC	ADLB. LBSTNRLO ADLB. LBSTNRHI
001-0001 51/F/White	DRUG A	Alanine Aminotransferase	Chemistry	Screening	2016-02-05	44/ U/L	No	6/ 34
		Albumin	Chemistry	Visit 8	2016-02-26	43/ g/L	Yes	33/ 49

III. SUMMARY-LEVEL CLINICAL SITE DATASET

a. Key Process Steps:

1. Programmer should refer SAP and Protocol of each studies to understand their primary efficacy endpoints.

2. Perform a reconciliation between the Site-level Investigator file (Standard Siteinfo.xls) and ADSL, SDTM DM datasets in order to check for inconsistencies like duplicate records, and any other missing needed information. If discrepancies are found, the trial sponsor must be contacted to resolve the issue. Convert the excel file to a SAS-dataset.

3. The following variables in Site-level Investigator file should have non-missing values: STUDYID, STUDYTL, SPONCNT, SPONNAME, NDA or BLA, SITEID, LASTNAME, FRSTNAME, PHONE, COUNTRY, CITY, STREET. If STREET variable exceeds 200 characters, an additional variable STREET1 can be created. It should be noted that variable values cannot exceed 200 characters.

4. Any missing data in Site-level Investigator file must be handled as per the instructions in current FDA's Bioresearch Monitoring Technical Conformance Guide. For example, if POSTAL value is missing or not applicable, keep the values as "NA".

5. For each study, create separate dataset specification and clinsite dataset by referencing the site level investigator file, SDTM and ADaM datasets.

6. While creating clinsite dataset in step 5 above, special attention is needed to merge the required SDTM datasets such as DM, DS, AE, DV and DD with ADSL or merge respective ADaM datasets. Sort by STUDYID and SITEID or by STUDYID, SITEID and ARM. Create required variables such as SAFPOP, SCREEN, DISCSTUD, DISCTRT, NSAE, SAE, DEATH, PROTVIOL, CENSOR based on BIMO conformance guide.

7. Based on Study endpoints, six efficacy variables are to be created for each study: ENDPTYPE, ENDPOINT, TRTEFFR, TRTEFFS, SITEEFFE, and SITEEFFS.

8. While programming, in parallel, start creating define.xml which is the metadata file that explains each variable and its derivation algorithm.

Note: Text in the "Derivation" column cannot exceed 1000 characters. If greater than 1000 characters, then add the text in a pdf file (longderivation.pdf) and provide the link. The pdf file should be hyperlinked to define.xml.

9. The created dataset is merged with the Site-Investigator dataset using STUDYID and SITEID.

10. All the individual study datasets are stacked together in order to create the final submission ready "clinsite" dataset. The structure of this dataset should conform to FDA's requirement of having one record per STUDY, per SITE, per ARM and at least the 39 variables which are necessary as per the BIMO conformance guide.

11. An independent validation of the specification and clinsite dataset must be performed in order to ensure they are complete and fully compliant as per the BIMO conformance guide.

12. To include in the submission, the clinsite dataset must be converted to SAS® transport file "clinsite.xpt".

Site-level Investigator file Template:

As per our work process, at the initial stage of BIMO work, the necessary information from each site is collected from the client using the Siteinfo.xls template. Table 3 represents a screenshot of this file.

Table 3. Screenshot of Site-level Investigator file to collect Site Level investigator details.

STUDYID	STUDYTL	SPONCNT	SPONNAME	IND	UNDERIND	NDA	BLA
Study Identifier	Study Title	Sponsor Count	Sponsor Name	IND Number	Under IND	NDA Number	BLA Number
ABCD203	A Phase 3, Randomized, Double-Blind Study to Compare the Efficacy and Safety	1	PharmaCo, Inc.	010020	Y	200056	

SUPPNUM	SITEID	FINLDISC	LASTNAME	FRSTNAME	MINITIAL	PHONE	FAX
Supplement Number	Study Site Identifier	Financial Disclosure Amount	Investigator Last Name	Investigator First Name	Investigator Middle Initial	Investigator Phone Number	Investigator Fax Number
4	007	$25,000	xxx	xxx	xx	xxx-xxx-xxxx	xxx-xxx-xxxx

EMAIL	COUNTRY	STATE	CITY	POSTAL	STREET	STREET1
Investigator Email Address	Country	State	City	Postal Code	Street Address	Street Address Continued
abcd@gmail.com	USA	PA	Berwyn	xxx	xxx	xxx

Note: Any variable with a length greater than 200 characters, must be split to more variables.
Row1: SAS Variable Name; Row2: Variable Label; Row3: example

b. Handling Efficacy Variables:

Deriving values for efficacy variables is a challenging piece in clinsite dataset. One has to clearly understand the primary efficacy endpoints of each study and select appropriate variables/values from the respective ADaM datasets (Eg: ADTL, ADEFF). The efficacy results are further summarized based on the category of data.

Few case studies for creating efficacy variables:
Case Study 1:

Efficacy Variable	Derivation
ENDPOINT	The primary efficacy analysis is a responder analysis comparing the two treatment groups based on the proportion of subjects with all Target Lesions are Clear on the Physician Lesion Assessment (PLA=Clear) at Visit 8 (Day 106).
ENDTYPE	Categorical
Treatment Efficacy Result (TRTEFFR)	If ADSL.STUDYID in (A-101, A-103) and ADTL.PARAMCD="LESION" and ADTL.ARM not missing and ADTL.AVISITN = 106 then check for Face PLA location to be Clear. This will be the number of Treatment success. For Proportion of Treatment success at each site and treatment get the TRTEFFR as No. of subjects with Treatment success/no. of subjects present with the treatment in the site.
Treatment Efficacy Result Standard Deviation (TRTEFFS)	The TRTEFFS is determined as (TRTEFFR) *(1-TRTEFFR)/no. of subjects present with the treatment in the site.
Site-Specific Treatment Effect (SITEFFE)	For ADSL.STUDYID in (A-101, A-103) and ADTL.PARAMCD="LESION" and ADTL.ARM ne '' '' and ADTL.AVISITN = 106 then check for Face PLA location to be Clear. This will be the number of Treatment success. For Proportion of Treatment success at each site get the SITEEFFE as No. of subjects with Treatment success/no. of subjects present in the site.
Site-Specific Treatment Effect Std (SITEEFFS)	The SITEEFFS is determined as (SITEEFFE) *(1-SITEEFFE)/no. of subjects present in the site.

Sample Data:

USUBJID	SITEID	PARAMCD	AVISIT	AVISITN	ADY	AVALC	ARM	EVENT	ENROLL
A-101-001-0001	1	LESION	Visit 8	106	103	THICK	PLACEBO	Y	Y
A-101-001-0002	1	LESION	Visit 8	106	103	THIN	DRUG A	N	Y
A-101-001-0003	1	LESION	Visit 8	106	107	CLEAR	DRUG A	Y	Y
A-101-002-0001	2	LESION	Visit 8	106	107	THIN	DRUG A	N	Y
A-101-002-0002	2	LESION	Visit 8	106	103	THIN	DRUG A	Y	Y
A-101-002-0003	2	LESION	Visit 8	106	103	THICK	PLACEBO	Y	Y

USUBJID	SITEID	PARAMCD	AVISIT	AVISITN	ADY	AVALC	ARM	EVENT	ENROLL
A-101-003-0001	3	LESION	Visit 8	106	106	CLEAR	DRUG A	Y	Y
A-101-003-0002	3	LESION	Visit 8	106	106	THICK	PLACEBO	Y	Y
A-101-003-0003	3	LESION	Visit 8	106	106	THIN	DRUG A	N	Y
A-101-003-0004	3	LESION	Visit 8	106	104	CLEAR	DRUG A	Y	Y

Sample code:

```
Proc freq data=ADTL noprint;
    where Event="Y";
    table siteid*event*arm / nopercent out=STAT1 (keep=SITEID ARM COUNT);
run;

Proc sort data=STAT1 (Rename=(COUNT=TRTSUC)) out=STAT2;
    by siteid arm;
run;

proc sort data=enroll;
    by siteid arm;
run;

data trteff;
    merge enroll stat2;
    by siteid arm;
    if n(trtsuc, enroll)=2 then TRTEFFE=trtsuc/enroll;
    if trteffe ne . then trteffs1=((trteffe)*(1-trteffe))/enroll;
    else trteffs1=0;
    if trteffe=. then trteffe=0;
    if trteffe ne . then trteffr11=strip(put(trteffe,8.2));
    trteffr=input(trteffr11,8.2);
    if trteffs1 ne . then trteffs11=strip(put(trteffs1,8.3));
    trteffs=input(trteffs11,8.3);
    keep siteid arm trteffr trteffs;
run;


Proc freq data=ADTL noprint;
    where Event="Y";
    table siteid*event / nopercent out=STAT3 (keep=SITEID COUNT);
run;

Proc sort data=STAT3 (Rename=(COUNT=TRTSUC)) out=STAT4;
    by siteid;
run;

proc sort data=enroll;
    by siteid;
run;
```

```
data siteeff;
    merge screen stat4;
      by siteid;
      if n(trtsuc,screen)=2 then siteeffe1=trtsuc/enroll;
      if siteeffe1 ne . then trteffs1=((siteeffe1)*(1-siteeffe1))/enroll;
      else trteffs1=0;
      if siteeffe1=. then siteeffe1=0;
      if siteeffe1 ne . then siteeffe11=strip(put(siteeffe1,8.3));
      siteeffe=input(siteeffe11,8.3);
      if trteffs1 ne . then trteffs11=strip(put(trteffs1,8.4));
      siteeffs=input(trteffs11,8.4);
      keep siteid siteeffe siteeffs;
run;
```

Case Study 2:

Efficacy Variable	Derivation
ENDPOINT	Percent change from baseline in seizure count
ENDTYPE	Continuous
Treatment Efficacy Result (TRTEFFR)	If ADSZ.PARAMCD in ("SEIZURCT") and ADSZ.BASE not equal 0 then PCHG is derived as ((AVAL−BASE)/BASE)∗100. If ADSZ.BASE = 0 and ADSZ.AVAL = 0 then PCHG is set to 0 then calculate the Mean of PCHG grouped by STUDYID, SITEID, TRT01P
Treatment Efficacy Result Standard Deviation (TRTEFFS)	If ADSZ.PARAM in ("SEIZURCT ") and ADSZ.BASE not equal 0 then PCHG is derived as ((AVAL−BASE)/BASE)∗100. If ADSZ.BASE = 0 and ADSZ.AVAL = 0 then PCHG is set to 0 then calculate the Standard Deviation of PCHG grouped by STUDYID, SITEID, TRT01P.
Site-Specific Treatment Effect (SITEFFE)	If ADSZ.PARAM in ("SEIZURCT ") and ADSZ.BASE not equal 0 then PCHG is derived as ((AVAL−BASE)/BASE)∗100. If ADSZ.BASE = 0 and ADSZ.AVAL = 0 then PCHG is set to 0 then calculate the Mean of PCHG grouped by STUDYID, SITEID.
Site-Specific Treatment Effect Std (SITEEFFS)	If ADSZ.PARAM in ("SEIZURCT ") and ADSZ.BASE not equal 0 then PCHG is derived as ((AVAL−BASE)/BASE)∗100 If ADSZ.BASE = 0 and ADSZ.AVAL = 0 then PCHG is set to 0 then calculate the Standard Deviation of PCHG grouped by STUDYID, SITEID.

Sample Data:

USUBJID	SITEID	PARAMCD	AVISIT	AVISITN	ADY	AVAL	BASE	CHG	ARM	ITTFL
B-001-0001	001	SEIZURCT	Visit 7	7	28	9	7	2	PLACEBO	Y
B-001-0002	001	SEIZURCT	Visit 7	7	28	120	50	70	DRUG A	Y
B-001-0005	001	SEIZURCT	Visit 7	7	28	2	6	-4	DRUG A	Y
B-002-0001	002	SEIZURCT	Visit 7	7	28	15	17	-2	PLACEBO	Y
B-002-0002	002	SEIZURCT	Visit 7	7	28	6	27	-21	DRUG A	Y
B-002-0004	002	SEIZURCT	Visit 7	7	28	3	8	-5	DRUG A	Y
B-003-0001	003	SEIZURCT	Visit 7	7	28	14	10	4	PLACEBO	Y

USUBJID	SITEID	PARAMCD	AVISIT	AVISITN	ADY	AVAL	BASE	CHG	ARM	ITTFL
B-003-0006	003	SEIZURCT	Visit 7	7	28	13	7	6	PLACEBO	Y
B-003-0003	003	SEIZURCT	Visit 7	7	28	16	24	-8	DRUG A	Y
B-003-0004	003	SEIZURCT	Visit 7	7	28	30	35	-5	DRUG A	Y

Sample Code:

```
Data eff;
     set adsz;
     where paramcd in ("SEIZURCT") and aval ne . and ITTFL="Y";
     if base ne 0 and aval ne . then pchg = ((aval-base)/base)*100;
     else pchg = . ;
     if base = 0 and aval =0 then pchg =0;
     keep studyid usubjid siteid arm aval base pchg;
run;

proc sort data=eff;
    by studyid siteid arm;
run;

***********STUDYID SITEID ARM*********************;

proc means data=eff noprint ;
    by studyid siteid arm;
    var pchg;
    output out=armeff
    mean=TRTEFFR std=TRTEFFS;
run;

***********STUDYID SITEID*********************;

proc means data=eff noprint;
    class studyid siteid ;
    var pchg;
    output out=siteff
    mean=SITEEFFE std=SITEEFFS;
run;
```

c. Handling Missing Information in Clinsite dataset:

There are possibilities of missing information. Example: a site with screen failure subjects will have missing information for all derived variables. Similarly, a site has treated arm information but there are no subjects related to Death, Discontinuation or SAE. It is recommended to keep “0” for all missing, non-efficacy, numeric derived variables. The missing values of efficacy variables should not be imputed.

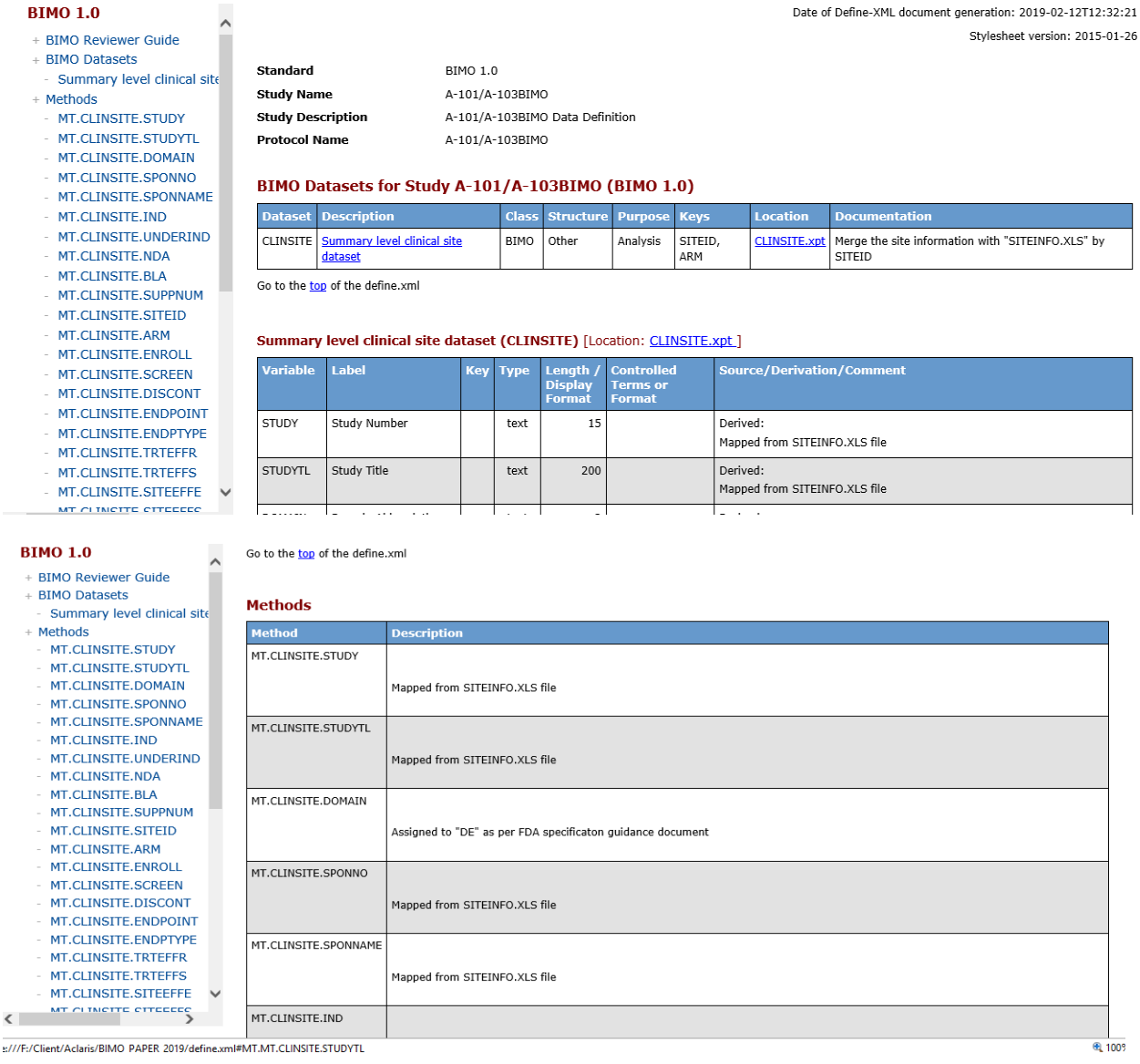
IV. DEFINE.XML

Define document otherwise called as metadata is a supportive document which contains all the necessary information about the site-level summary dataset and its location; attributes of each variable and the algorithm applied to derive it. As per Bioresearch Monitoring Technical Conformance Guide, clinsite dataset must contains at least 38 variables. This document helps the FDA reviewer to understand and correlate the data from the dataset with their analysis reports.

a. Key Process Steps:

1. Create the clinsite data specification excel file including the following columns: Variable, Label, Key, Type, Length and Description.
2. Add details of each one of 38 required variables as per the Conformance guide.
3. Make sure that the number of variables and their attributes in the site-level summary dataset agrees with clinsite data specification excel.
4. Perform QC checks to ensure completeness of the details provided.
5. Use clinsite specification excel file and site-level summary dataset to create a define document in PDF format through XML application.
6. Review the define PDF file and preform quality checks. The file should contain all the information from the site-level summary dataset.

b. Sample Define file:



V. BIMO REVIEWER'S GUIDE

As per BIMO conformance guide, Table 4 provides the BIMO Reviewer's Guide template. The guide contains a description of the BIMO elements that were requested with hyperlinks to those elements in Module 5.

Table 4: BIMO Reviewer's Guide template – our experience with FDA submissions supporting multiple therapeutic areas for several clients.

Section I Item 1	Please include the following information in a tabular format in the original NDA for each of the completed pivotal clinical trials: a. Site number b. Principal investigator c. Site Location: Address (e.g., Street, City, State, Country) and contact information (i.e., phone, fax, email) d. Location of Principal Investigator: Address (e.g., Street, City, State, and Country) and contact information (i.e., phone, fax, email). If the Applicant is aware of changes to a clinical investigator's site address or contact information since the time of the clinical investigator's participation in the study, we request that this updated information also be provided.	A list and description of investigators (with contact information) can be found in A-101 CSR Appendix 16.1.4 A-103 CSR Appendix 16.1.4 OR Hyperlink the pdf file Table 1
Section I Item 2	Please include the following information in a tabular format, by site, in the original NDA for each of the completed pivotal clinical trials: a. Number of subjects screened at each site b. Number of subjects randomized at each site c. Number of subjects treated who prematurely discontinued for each site by site	A-101 CSR Appendix 16.1.4 A-103 CSR Appendix 16.1.4 OR Hyperlink the pdf file Table 2
Section I Item 3	Please include the following information in a tabular format in the NDA for each of the completed pivotal clinical trials: a. Location at which sponsor trial documentation is maintained (e.g., monitoring plans and reports, training records, data management plans, drug accountability records, IND safety reports, or other sponsor records as described ICH E6, Section 8). This is the actual physical site(s) where documents are maintained and would be available for inspection. b. Name, address and contact information of all Contract Research Organization (CROs) used in the conduct of the clinical trials and brief statement of trial related functions transferred to them. If this information has	A-101 CSR Appendix 16.1.13 A-103 CSR Appendix 16.1.13 b.List of Vendor information can be directly provided here. OR Provide details in a PDF file and hyperlink here.

	been submitted in eCTD format previously (e.g. as an addendum to a Form FDA 1571, you may identify the location(s) and/or provide link(s) to information previously provided. c. The location at which trial documentation and records generated by the CROs with respect to their roles and responsibilities in conduct of respective studies is maintained. As above, this is the actual physical site where documents would be available for inspection.	c. Trial documentations and records are available at <SPONSER>, as part of the Trial Master Files (TMF). Some additional CRO generated documentation, not part of the TMF, are maintained at the CRO locations. Contact information for the CROs, can be found in Section 1, Item 3b.
Section I Item 4	For each pivotal trial, provide a sample annotated Case Report Form (or identify the location and/or provide a link if provided elsewhere in the submission).	A-101 Annotated CRF A-103 Annotated CRF
Section I Item 5	For each pivotal trial provide original protocol and all amendments (or identify the location and/or provide a link if provided elsewhere in the submission).	A-101 Protocol (02-Jun-2015) A-103 Protocol(04-Jun-2015)
Section II Request for Subject Level Data Listings by Site Item 1	For each pivotal trial: Site-specific individual subject data listings (hereafter referred to as "line listings"). For each site, provide line listings for: a. Listing for each subject consented/enrolled; for subjects who were not randomized to treatment and/or treated with study therapy, include reason not randomized and/or treated b. Subject listing for treatment assignment (randomization) c. Listing of subjects that discontinued from study treatment and subjects that discontinued from the study completely (i.e., withdrew consent) with date and reasondiscontinued d. Listing of per protocol subjects/ non-per protocol subjects and reason not per protocol e. By subject listing of eligibility determination (i.e., inclusion and exclusion criteria) f. By subject listing, of AEs, SAEs, deaths and dates g. By subject listing of protocol violations and/or deviations reported in theNDA, including a description of the deviation/violation h. By subject listing of the primary and secondary endpoint efficacy parameters or events. For derived or calculated endpoints, provide the	Study BIMO Listings - A-101 001, 002, 003, 004, 005 Study BIMO Listings - A-103: 001, 002, 003, 004, 005

	raw data listings used to generate the derived/calculated endpoint. i. By subject listing of concomitant medications (as appropriate to the pivotal clinical trials) j. By subject listing, of testing (e.g., laboratory, ECG) performed for safety monitoring	
Section III Request for Site Level Dataset	Site level datasets	Study BIMO - Data Definitions Table - XML Study BIMO - clinsite - Summary Level Clinical Site Dataset

Table 1:

Please include the following information in a tabular format, by site; Site number, Principal investigator, Site Location: Address and contact information.

Site ID	Principal Investigator	Site Location: Address and Contact Information
101	xxx	Dermatology Division xxx, Suite xxx, xxx, xxx, xxx, Postal: xxxxx, Email: xxxxx@gamil.com Phone: xxx-xxx-xxx; Fax: xxx-xxx-xxx

Table 2:

Please include the following information in a tabular format, by site; Number of subjects screened at each site, Number of subjects randomized at each site, Number of subjects treated who prematurely discontinued for each site by site.

Study ID	Site ID	Number of Subjects Screened	Number of Randomized Subjects	Number of Subjects Treated Who Prematurely Discontinued
A-101	101	10	7	2

CREATING THE DATA FILE (TEMPLATE AND STRUCTURE)

- Construct a BIMO study tagging file (STF) and place it in eCTD Module 5.3.5.4. 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
Bimo Reviewer's guide	Site Investigator details	.pdf

CONCLUSION

BIMO continues to be one of the critical requirements of the regulatory agencies to assess the quality and integrity of the study data at investigator sites during a clinical trial and provide feedback to the sponsor. The clinical programmer access clinical trial data from the database and generates necessary datasets and documents. Programmers must understand the regulatory requirements; the endpoints of each study and the completeness of Site-level investigator file. In case of multiple studies, programmer needs to combine multiple studies information into a single submission dataset and a document as outlined in the Bioresearch monitoring technical conformance guide. The authors have applied their working experience to generate this paper which can serve as a comprehensive process guideline to develop quality BIMO submission documents that would increase the probability of acceptance from regulatory agencies and speedup approval process.

REFERENCES

1. Charanjit Singh Kahlon, Dharmendra Tirumalasetti, Bhavin Busa, and Kristie Kooken (2018) Programmer's Guide for OSI Deliverables – Creation of Site Level Summary Dataset and Automation of BIMO Listings Generation. PharmaSUG- Paper #SS16.

2. Ellen Lin, Wei Cui, Ran Li, and Yaling Teng (2018) Preparing the Office of Scientific Investigations (OSI) Requests for Submissions to FDA. PharmaSUG- Paper #EP15.

3. Bioresearch Monitoring Technical Conformance guide (2018). U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). https://www.fda.gov/downloads/drugs/developmentapprovalprocess/formssubmissionrequirements/ucm332468.pdf

4. Standardized Format for Electronic Submission of NDA and BLA Content for the Planning of Bioresearch Monitoring (BIMO) Inspections for CDER Submissions. https://www.fda.gov/downloads/drugs/developmentapprovalprocess/formssubmissionrequirements/ucm332466.pdf

ACKNOWLEDGMENTS

We would like to thank our colleagues Alfredo Rojas, Thomas Mannering and Latha Donapati for their valuable suggestions during preparation of this paper. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. Other brand and product names are trademarks of their respective companies.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

1. Ramanjulu Valluru
Accenture Inc,
1160 W. Swedesford Road Bldg. One
Berwyn, PA 19312
Phone: +1 (484) 690-5182
ramanjulu.valluru@accenture.com

2. Harsha H. Dyavappa
Senior Statistical Programmer,
Accenture Inc,
Bangalore, India.
Phone: +91 9980867867
harsha.h.dyavappa@accenture.com