

## Programmer’s Guide for OSI Deliverables – Creation of Site Level Summary Dataset and Automation of BIMO Listings Generation

Charanjit Singh Kahlon, Vita Data Sciences, Waltham, MA, USA  
 Dharmendra Tirumalasetti, Vita Data Sciences, Waltham, MA, USA  
 Bhavin Busa, Vita Data Sciences, Waltham, MA, USA  
 Kristie Kookan, Achaogen, San Francisco, USA

### ABSTRACT

The FDA conducts site inspections as part of the regulatory review process to ensure clinical trial is conducted according to the FDA regulations. For the drug related application (i.e. NDA), site inspection is managed by Office of Scientific Investigations (OSI) within CDER office of compliance. The OSI requests that the Sponsor should submit data that describes the characteristics and outcomes of clinical investigations at the site level as part of the NDA. The part II of the OSI request, which is mandatory, is to submit subject-level data listings by site (“BIMO listings”) for all major trials used to support safety and efficacy in the application. The requested BIMO listings for the most part is standard from one application to the other. The part III of the OSI request, which is optional, is to submit summary level clinical site dataset (‘clinsite’). There is an FDA draft guidance and specifications document for preparing and submitting the dataset. Although the specifications provide details about the variables, the description provided for efficacy variables can be expanded further, considering primary end points may vary for each application. This paper provides details about each variable and how to derive it from various data sources. Special emphasis is given on the efficacy variables, explaining different formulae used for different types of primary end points. i.e. continuous, discrete, etc. This paper also includes SAS macro to automate the generation of BIMO listings in pdf with bookmarks for each site, using the SDTM/ADAM datasets, per the OSI requirements.

### INTRODUCTION

The current format of subject level data in NDA and BLA submission packages from sponsors may not be sufficient for site selection process. In order to provide a robust method for selecting sites for inspections, OSI requests sponsor to submit general study related and investigator information, subject level data listing by site for pivotal studies, and site level dataset created following FDA guidelines<sup>[2]</sup>. The FDA Bioresearch Monitoring Program (BIMO) develops guidelines for inspections of clinical investigators, sponsors, and Institutional Review Boards (IRBs). The FDA Office of Scientific Investigations (OSI) manages the BIMO program for drugs, and the FDA Office of Inspections and Surveillance manages the BIMO program for biologics.

There are three parts to the BIMO request to be provided to the FDA: part I includes general study related information and specific Clinical Investigator information, part II is delivering subject-level data listings for pivotal studies and part III is voluntary and focuses on developing a site-level dataset in a standardized electronic format<sup>[2]</sup>.

### PART 1: GENERAL STUDY RELATED INFORMATION

#### A. Comprehensive and Readily Located List of All Clinical Sites<sup>[2, 3]</sup>

Below is recommended format for comprehensive and readily located list of all clinical sites which participated in the pivotal studies.

Study Identifier	Site Identifier	Investigator Name	Site Address at Time of Clinical Study	Site Contact Information at Time of Clinical Study
STUDYID	SITEID	LASTNAME, FRSTNAME, INITIAL	FACILITY NAME STREET CITY, STATE, POSTAL COUNTRY	PHONE FAX EMAIL

XXXX	XXXX	XXXX, XXXX, X.	xxxxxxxxxxxxxxxxxxx	<b>Phone: 1-555-555-5555</b> <b>Fax: 1-555-555-5555</b> <b>Email: xxx.x@mail.com</b>
------	------	----------------	---------------------	--

**B. Table Listings of All Entities to Whom Sponsor Has Contracted Clinical Study-Related Activities** <sup>[2, 3]</sup>

The FDA guidelines for BIMO expect a tabular format of listing entities that include contracted clinical study-related activities in PDF form for pivotal studies during the submission process, location of study-related documents for each study provided by sponsors, and who had generated it (i.e. whether sponsor or CRO). These documents include monitoring plans and reports, training records, and data analysis plans (Trial Master File documents). The sponsor also provides contact information (i.e., phone number and email address) for the individual(s) who can provide updated location of study documents to FDA if needed. This information ensures the records are accurate for review once the inspection is assigned.

**C. Protocol, Protocol Amendments, and Annotated Case Report Form** <sup>[2, 3]</sup>

This section includes the protocol, protocol amendments, and case report form in Appendix 16 of clinical study report for all studies. If sponsor is going to submit the BIMO Reviewer’s Guide, it can be included in Appendix 16 along with the above information. Hyperlinks to their locations in the submission should be included. These items are required by the Office of Regulatory Affairs for BIMO inspections and it is important to submit all versions of these documents so that field inspectors performing site inspection refer to the correct documents.

**PART II: SUBJECT LEVEL DATA LISTINGS BY SITE**

The subject-level data listings should be submitted by site. Sites that are involved in multiple studies in support of an application should submit these listings separately for each study within the submission. These listings should include below information:

- a. Listing of screened subject and reason for subjects who did not meet eligibility requirements
- b. Subject listing for treatment assignment (randomization)
- c. Listing of subject discontinuations with date and reason
- d. Listing of subject’s evaluability
- e. Listing of eligibility determination
- f. Listing of adverse events and deaths
- g. Listing of important protocol violations/deviations
- h. Listing of efficacy parameters
- i. Listing of concomitant medications
- j. Listing of tests performed for safety monitoring

The mock-shell for all listings are presented below. The content of these mocks is prepared based on our experience and may need slight modifications based on the study. The reader is recommended to contact OSI for any clarification while working on this request. The authors’ intention is to provide an idea how these listings might look like and not to propose standard templates.

Source data for these listings should be the same as the individual clinical study reports (i.e., SDTM or ADaM datasets). All listings should be developed by site by subject. Typically, the site identification is included as part of the output with subject identifiers and other relevant information as columns and listed chronologically.

**a. Listing of screened subjects and reason for subjects who did not meet eligibility requirements**

This listing should include all consented subjects who enrolled. Subjects who consented but were screen failures should also be included. Additionally, the reason for those subjects who consented but not randomized or treated is included in this listing

Subject Identifier	Randomized Population	Treated Population	Reason not Eligible
XXX-XXX-XX	X	X	XXXXXXXXXXXXXXXXXXXX
XXX-XXX-XX	X	X	XXXXXXXXXXXXXXXXXXXX

**b. Subject listing for treatment assignment (randomization)**

This listing should include the treatment of assignment to which subject was randomized. If the treatment assignment to randomized subjects differed from treatment subjects were received, include actual treatment as well.

Subject Identifier	Treatment Assignment	Treatment Received
XXX-XXX-XX	XXXX	XXXX
XXX-XXX-XX	XXXX	XXXX

**c. Listing of Subject Discontinuations**

This listing is about discontinuations and should include all subjects who discontinued from treatment/study. For studies with run-in periods, the subjects discontinued during run-in-period should also be provided. The date and reason of discontinuations should also be provided for each subject within a site.

Subject Identifier	Treatment Assignment	Reason of Study Discontinuation	Date of Discontinuation
XXX-XXX-XX	XXXX	XXXXXXXXXXXXXXXX	XXXX-XX-XX
XXX-XXX-XX	XXXX	XXXXXXXXXXXXXXXX	XXXX-XX-XX

**d. Listing of Subject Evaluability**

This listing should include the subject population as defined in the protocol (e.g., intent-to-treat, safety, microbiological, per-protocol etc.) as well as any subjects that are not part of these populations. These population flags are mostly derived in analysis datasets or provided in supplemental SDTM datasets.

Subject Identifier	ITT Evaluable	Reason Not Evaluable for ITT	Safety Evaluable	Reason Not Evaluable for Safety	Per-Protocol Evaluable	Reason Not Evaluable for Per-Protocol
XXX-XXX-XX	X	XXXXXXX	X	XXXXXXX	X	XXXXXXX
XXX-XXX-XX	X	XXXXXXX	X	XXXXXXX	X	XXXXXXX

**e. Listing of Eligibility Determination**

This listing should include whether inclusion/exclusion were met as defined in protocol. The listing should provide the list of the inclusion/exclusion criteria violated.

Subject Identifier	Met All Incl/Excl Criteria	Inc/Excl Criteria Violated
XXX-XXX-XX	X	XXXXXXXXXXXXXXXXXXXX
XXX-XXX-XX	X	XXXXXXXXXXXXXXXXXXXX

**f. Listing of Adverse Events and Deaths**

This listing should include all adverse events, the adverse event onset and resolution date, whether adverse event is serious, severity of adverse events, action taken for any adverse events and outcome for adverse events.

Subject Identifier	Preferred Term	Onset Date/Resolution Date	Serious	Death due to AE	Action Taken	Outcome of AE
XXX-XXX-XX	XXXXXXXX	XXXX-XX-XX /XXXX-XX-XX	X	X	XXXXX	XXXXXXXXXX
XXX-XXX-XX	XXXXXXXX	XXXX-XX-XX /XXXX-XX-XX	X	X	XXXXX	XXXXXXXXXX

**g. Listing of Protocol Violations/Deviations**

This listing should include all major protocol deviations. The description of each protocol deviation should also be provided.

Subject Identifier	Treatment Assignment	Deviation	Category	Protocol Violation/Deviations
XXX-XXX-XX	XXXX	XXXXXXXX	XXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXX-XXX-XX	XXXX	XXXXXXXX	XXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX

**h. Listing of Efficacy Parameters**

This listing should contain primary and main secondary efficacy parameters. For derived or calculated endpoints, the source data points (from SDTM, ADaM or other if another source exists) used to derive primary and secondary endpoints should also be provided.

Subject Identifier	Primary Endpoints	Secondary Endpoints
XXX-XXX-XX	XXX	XXX
XXX-XXX-XX	XXX	XXX

**i. Concomitant Medications**

This listing includes all concomitant medications that are collected per protocol. This also includes start/end dates of concomitant medication, name of medication, dose, route and reason for administration.

Subject Identifier	Treatment Received	Indication	Medication Term	Start Medication /End Medication	Dose/Unit/Route	Reason for Administration
XXX-XXX-XX	XXXX	XXXXXXXX X	XXXXXX XX	XXXX-XX-XX / XXXX-XX-XX	XX/XX/XXXX	XXXXXXXXXXXX
XXX-XXX-XX	XXXX	XXXXXXXX X	XXXXXX XX	XXXX-XX-XX / XXXX-XX-XX	XX/XX/XXXX	XXXXXXXXXXXX

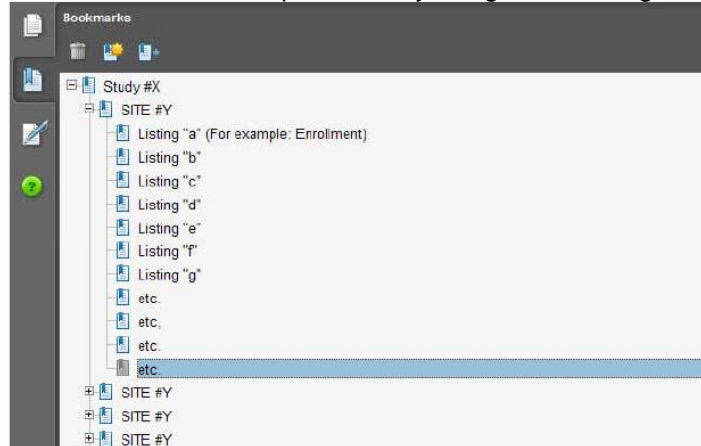
**j. Safety Monitoring**

This listing includes the test performed as part of safety monitoring. It mainly includes finding domains such as LB, EG etc. The information collected for safety monitoring is as defined in the protocol.

Subject Identifier	Treatment Assignment	Category	Test Name (Unit)	Visit	Collection Date/Time	Result
XXX-XXX-XX	XXXX	XXXXXXXX	XXXXXXXX	XXXX	XXXX-XX-XX	XX
XXX-XXX-XX	XXXX	XXXXXXXX	XXXXXXXX	XXXX	XXXX-XX-XX	XX

## AUTOMATE GENERATION OF SUBJECT LEVEL DATA LISTINGS (PART II) WITH BOOKMARKS

The generation of subject level data listings is a laborious task and pre-planning and having a process in place is important to meet this important NDA deliverable. It is expected that the Sponsor provides BIMO data listings from all major (e.g. pivotal) studies supporting safety and efficacy in the application. FDA requests that one PDF file be created for each pivotal study using the following format: <sup>[1]</sup>



The expectation is that the bookmark is created for each listing and for every site. Based on our experience, it is easy to assign bookmarks in RTF file. The expected format for submission of BIMO is PDF. The bookmarks from RTF file may not work as expected once it is converted to PDF. It is also very tedious to work with the PDF file to set-up bookmarks as expected. We worked on this issue and found alternative way to address it. The following section will explain the generation of BIMO listings with use of macro and automation of generating bookmarks using SAS procedure, PROC DOCUMENT.

Our program was based on the template of BIMO listings provided in the previous section. We created a dataset for each listing from SDTM & ADaM datasets and combined all of them into one single dataset for all sites. A variable 'listnum' has been added to represent each listing for all sites, which is also used to create macro variable for each listing out of all BIMO listings. We used our standard macro to create report for each listing in all sites with traditional ODS PDF and the following output has been generated. This is not in the desired format for bookmarks where none of the bookmarks are connected to a single top node (study) and site level node.

BIMO\_Listings.pdf - Adobe Acrobat Standard DC  
File Edit View Window Help

Home Tools Document 2 / 50

Bookmarks

- Study xxxx
  - Site 01
    - Listing a
    - Listing b
  - Site 01
    - Listing c
  - Site 01
    - Listing d
  - Site 01
    - Listing e
  - Site 01
    - Listing f

**b. Subject listing for treatment assignment (randomization)**

Unique Subject Identifier	Description of Planned Arm	Description of Actual Arm
SOFT-101-01001	COHORT 1 XXXX	COHORT 1 XXXX
SOFT-101-05001	COHORT 1 XXXX	COHORT 1 XXXX
SOFT-101-06001	COHORT 1 XXXX	COHORT 1 XXXX
SOFT-101-06002	COHORT 1 XXXX	COHORT 1 XXXX
SOFT-101-06003	COHORT 2 YYYY	COHORT 2 YYYY
SOFT-101-06004	COHORT 2 YYYY	COHORT 2 YYYY
SOFT-101-08002	COHORT 1 XXXX	COHORT 1 XXXX

**Programmer's Guide for OSI Deliverables – Creation of Site Level Summary Dataset and Automation of BIMO Listings Generation, Continued**

However, In SAS 9.2 and later versions, you can run PROC REPORT with ODS DOCUMENT and this new capability enables you to store the REPORT output which can be used to replay to different destinations (such as PDF) using PROC DOCUMENT.

If PROC REPORT code (as below) is run using ODS DOCUMENT, the report structure is saved in a document and any necessary changes can be made to it within PROC DOCUMENT.

```
ods listing close ;
ods document name=bimolist(write) ;
%macro runlist ;

%do x= 1 %to &sttot. ; ** Total number of unique sites **;
  %list_report(dsn=all_list, sitevar=siteid,
              grpbyn=1, grpby1=usubjid,
              listbyn=3, listby1=randpop, listby2=treatpop,
              listby3=reas,
              cont=%str(Listing a),
              listnum=1,
              title = %str(a. Listing of screened subject and reason
for subjects who did not meet eligibility requirements)
              ) ;

  %list_report(dsn=all_list, sitevar=siteid,
              grpbyn=1, grpby1=usubjid, grpby1cw= 2.5in ,
              listbyn=2, listby1=arm, listby2=actarm,
              listby1cw=2in,listby2cw=2in,
              cont=%str(Listing b),
              listnum=2,
              title = %str(b. Subject listing for treatment
assignment (randomization))
              ) ;
  *.....(Report code/macro for each BIMO Listing) ;
%end ;
%mend ;
%runlist ;
ods document close ;
```

Within the following code, MOVE statement in PROC DOCUMENT is used to rearrange the table of contents or structure and then can be replayed using ODS PDF:

```
%macro bimopdf ;
proc document name=bimolist;
  *** Move all the listing level nodes (listings a thru j) for a site from
  different reports under that site ***;
  %do z =1 %to &sttot. ;
    %let s=%eval(%eval(&z.-1)*10) ;
    %let r= 0 ;
    %do n=1 %to 10 ;
      %let r=%eval(&s. + &n.) ;
      move \Report#&r.\Report#1\Report#1 to \Report#&z.\Report#1;
    %end ;

    setlabel \Report#&z.\Report#1 "Site &st&z.";
  %end ;
  ** ** ;

  ** Move all the site level(with sub level) nodes under one single study
  node **;
  %do k=2 %to &sttot. ;
    move \Report#&k.\Report#1 to \Report#1 ;
  %end ;
```

## Programmer's Guide for OSI Deliverables – Creation of Site Level Summary Dataset and Automation of BIMO Listings Generation, Continued

```
** ** ;
setlabel \Report#1 "Study xxxx";
run ;

ods pdf file = "Z:\Study xxxx\BIMO\BIMO_Listings.pdf";
replay;
run;
quit;

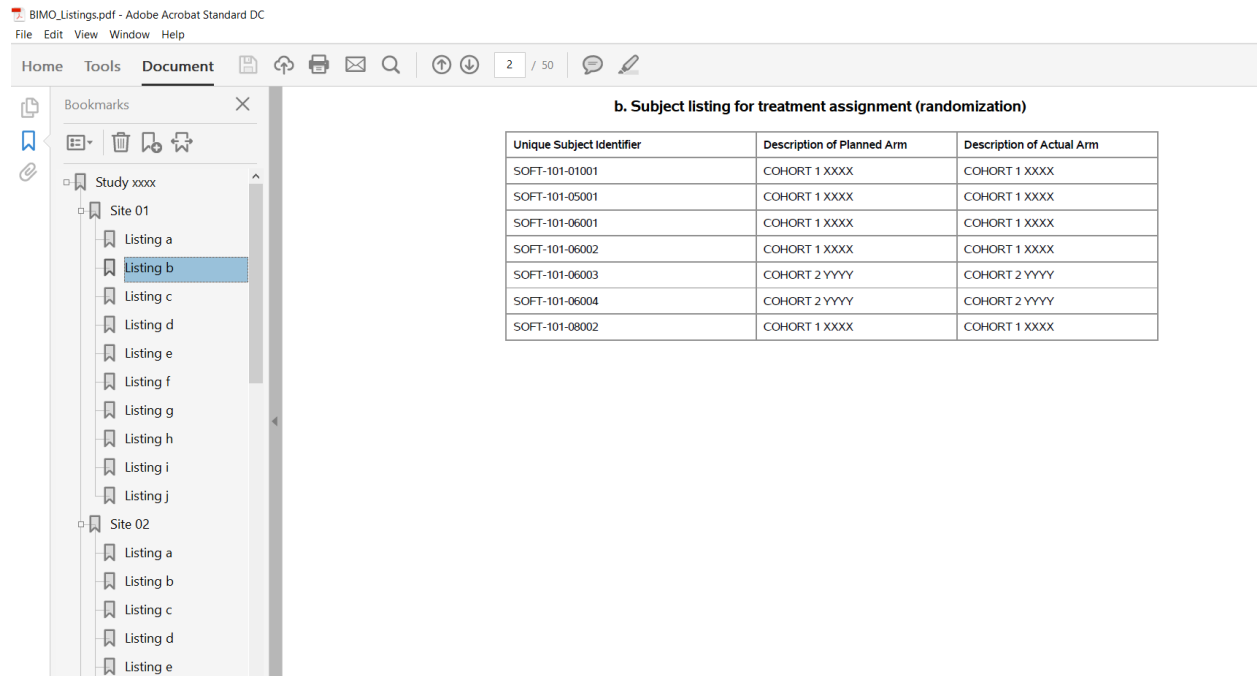
ods pdf close;
ods listing;

** To remove documents created - Make sure to run this immediately running the pdf **;
proc datasets lib=work kill memtype=(itemstor) ;
quit ;

** If need to rerun PDF part, make sure to run it from proc report part.**;
%mend ;

%bimopdf ;
```

After running the above code, A single PDF file with all the site-level BIMO listings is created keeping all the bookmarks intact as shown below:



### SITE LEVEL SUMMARY DATASET (PART III)

The site level summary dataset, which is optional to submit, is used by FDA, CDER for efficient site selection process. The FDA developed a risk-based tool that uses site level summary dataset to identify target sites for inspection. The FDA requests to submit a clinical site dataset for all pivotal studies involved in the NDA application as input to the risk-based tool. In addition, this dataset can also help Sponsor to identify any sites with potential issues in advance and prepare well for the inspection.

The clinical site dataset contains site level summaries for all pivotal studies used to support efficacy and safety in the submission, the studies with different treatment indications should also be provided. Clinical site data elements should include below variables as per the FDA guidance [2, 3].

### A. Study Level information

The study level information includes study identifier, study title, primary endpoint, primary endpoint type and description of planned treatment arm [2]. The information except treatment arm may be obtained from protocol or from TS domain and it should match with clinical study report or as present in the data definitions (define.xml).

### B. Submission Level Information

The submission level information includes sponsor count, sponsor name, IND number, under IND, NDA number, BLA number and supplement. The sponsor count variable should include all sponsors during the study, the sponsor name (full name or organization conducting the study at time of study completion) as defined in 21 CFR 312.3 [2]. If sponsor name is more than 200 characters, then provide short name that can be adjusted within 200 characters. The information IND number, under IND, NDA number, BLA number and supplement are study related information.

### C. Efficacy Information

As per FDA technical conformance guide [2]. The primary end points will have one of the 4 types of endpoints as explained in table 1.

Table 1: The primary endpoint types for different studies in clinical trials.

Primary Endpoint Type	Description
Discrete	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 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 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 (TRTEFFR) 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 value(s) with precisely defined variable interpretations should be submitted as part of the dataset.

In the past, we have created efficacy variables with discrete primary endpoints. For this paper, we will use one of the studies we have worked on as an example. The details of the efficacy variable derivations are provided in table 2.

Table 2: The primary endpoint types for various studies in clinical trials.

Efficacy Variables	Derivation
Treatment Efficacy Result (TRTEFFR)	The number of subjects responding to treatment ITT/SAFFL/other population in each treatment/ The number of subjects in ITT/SAFFL/other population within each treatment
Treatment Efficacy Result Standard Deviation (TRTEFFS)	Square root of $TRTEFFE*(1-TRTEFFE)/$ number of ITT/SAFFL/other subjects



Efficacy Variables	Derivation
Site-Specific Treatment Effect (SITEEFFE)	TRTEFFE for active treatment -TRTEFFE for placebo/comparative treatment. The variable is only populated for active drug records.
Site-Specific Treatment Effect Standard Deviation (SITEEFFS)	<p>Square root of <math>P*(1-P) *(1/n1+1/n2)</math>, where <math>P=(p1*n1+p2*n2)/(n1+n2)</math>.                      Where p1-The total number subjects with success for primary endpoint for active drug                      p2-The total number subjects with success for primary endpoint for placebo/comparative drug.</p> <p>n1=number of subjects randomized and treated in a site per active drug                      n2=number of subjects randomized and treated in a site per for placebo/comparative drug.</p>

#### D. Safety Information

The summary of safety information at site level includes number of subjects in safety population, number of non-serious adverse events, number of serious adverse events and number of deaths [2]. This information should be present in AE domain or any other domain. The counts should be calculated by sites and treatments.

#### E. Financial and site level Information

The financial and site level information includes study site identifier, number of subjects screened, number of subject discontinued from study, number of subjects discontinued from study treatment, number of protocol violations and financial disclosure amount [2]. All variables except financial disclosures can be sourced with study level SDTM domains. For financial disclosures, coordination with the Clinical Operations and Regulatory functions is key to determine the source to populate the financial fields.

#### F. Investigator Contact Information

These are the variables related to the contact details of investigators at different sites. The list includes Investigator Last Name, Investigator First Name, Investigator Middle Initial, Investigator Phone Number, Investigator Fax Number, Investigator Email Address, Country, State, City, Postal Code, Street Address, Street Address Continued [2]. This information is obtained from each site and received generally in a spreadsheet.

### CONCLUSION

The generation of subject-level data listings with bookmarks using SAS MACRO will be greatly helpful for submission. This could save time and resources that eventually could be spent on improving overall quality of submission package. The detailed explanation of variables of clinical site dataset and BIMO listings offer in depth understanding of these deliverables and is helpful while preparing FDA submission package.

### REFERENCE

1. Dreyer, T. and Scetinina, T. (2015) OSI Packages: What you need to Know for your next NDA or BLA Submission. PharmaSUG- Paper #SS05.
2. 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>

3. Standardized Format for Electronic Submission of NDA and BLA Content for the Planning of Bioresearch Monitoring (BIMO) Inspections for CDER Submissions Guidance for Industry (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/ucm332466.pdf>

## CONTACT INFORMATION

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

**Charanjit Singh Kahlon**

Organization: Vita Data Sciences, a division of Softworld, Inc

Address: 281 Winter St., Waltham, MA 02451

Work Phone: 781-466-8882

Email: [ckahlon@softworldinc.com](mailto:ckahlon@softworldinc.com)

Brand and product names are trademarks of their respective companies.