PharmaSUG 2022 - Paper QT-114

Avoid Common Mistakes in Preparing the BIMO Deliverables

Chintan Pandya, Merck & Co., Inc., Rahway, NJ, USA

Himanshu Patel, Merck & Co., Inc., Rahway, NJ, USA

ABSTRACT

The purpose of the BIMO package is to verify the integrity of the data submitted and confirm that all regulations for clinical trials are being met. The data package is required for pivotal studies and should include information to support your application's safety and efficacy endpoints. It is important to pay special attention when creating a BIMO package and avoid making mistakes. This paper will discuss common mistakes we encounter and avoid them to create a good quality summary-level clinical site (CLINSITE) dataset, subject-level line listings by Site, and data definition (define.xml). The paper also provides tips on cross-checking the efficacy and safety counts with CSR counts.

INTRODUCTION

The FDA's Bioresearch Monitoring (BIMO) programs require sponsors' help to select inspection sites quickly and efficiently, and the office of Scientific Investigations (OSI) has its responsibility. The BIMO package is divided into three parts as described below:

a) Clinical Study-Level Information

b) Subject-Level Data Line Listing by Site

c) Summary-Level Clinical Site (CLINSITE) dataset

The summary level clinical site dataset (clinsite.xpt) is one of the three components of the BIMO package and includes data from all pivotal studies. The CLINSITE dataset contains data from two sources, from ADaM/SDTM dataset and investigator contact information, which is made available by the clinical team in an excel format. Due to the multiple data source as an input, it is essential to carefully review the dataset program and dataset. The dataset is provided with the definition document (define.pdf) for review. We may face certain challenges when creating a CLINSITE dataset or define, and it can be avoided with proper tips and experience. This paper will discuss preventing common mistakes and creating good quality BIMO deliverables.

SUBJECT-LEVEL DATA LINE LISTING BY SITE

Subject-level data line listing by site is part II of the OSI request. There are 10 unique type listings to be generated by the clinical site, which serve as background data for inspectors to refer to during inspection. The sponsor should discuss with the Office of Scientific investigations (OSI) if they feel that additional listings or formats are required to support the application. The requirements are straightforward, and most companies have standard macros to generate these listings. However, it is important to understand the listing requirements (e.g., population use, data to display, delivery format, etc.).

The following are important checks to consider while generating subject-level data listing by site:

REQUIRED DATA FOR THE LISTING

The following table identifies the subject-level data line listing by the site and quick tips to ensure the required population and data selected when generating the listing.

Listing Type	Description
Consented Subjects (All Screened Subjects)	This listing should include all consented subjects, including screen failure. The reason for the subjects who consented but were not treated or randomized (screen failure) should be included. **Listing details example:** 1. Unique subject ID 2. Subject ID 3. Randomized (Y or N) 4. Treated (Y or N) 5. Reason, if screen failure
Treatment Assignment (All Randomized Subjects)	The listing should include the treatment assignment to which the subject was randomized and the actual treatment received by the subject. **Listing details example:** 1. Subject IDs 2. Planned Treatment 3. Actual Treatment
Discontinuation (All Randomized Subjects)	The listing should include all subjects discontinued from study treatment and from the study completely. **Listing details example:** 1. Subject ID 2. Treatment 3. Discontinuation date 4. Relative day of discontinuation 5. Reason for discontinuation 6. Type of discontinuation (e.g., treatment or study)
Study Population (All Randomized Subjects)	The listing identifies the study population in which each subject is analyzed. **Listing details example:** 1. Subject ID 2. Unique Subject ID 3. Treatment Group 4. ITT Population (Y or N) 5. ASaT Population (Y or N) 6. ApaT Population (Y or N) 7. Any Protocol-specified Population (Y or N)
Inclusion and Exclusion Criteria (All Screened Subjects)	The listing should display for each subject who did not meet the inclusion criteria or does meet the exclusion criteria. **Listing details example:** 1. Unique Subject ID 2. Subject ID 3. Reason (Inclusion or Exclusion Criteria)

Adverse Events (All Treated Subjects)	The listing should include all the adverse events (e.g., serious adverse events, non-serious adverse events, deaths due to adverse events, etc.), including other related information described below. **Listing details example:** 1. Subject ID 2. Treatment Group 3. AE Start Date / Day 4. AE End Date / Day 5. AE Preferred Term 6. AE Toxicity Grade 7. Serious (Y or N) 8. Related (Y or N) 9. Action Taken 10. AE Outcome (Display Death Date if the outcome is Death)
Important Protocol Deviations (All Important Protocol Deviations)	This listing should include all the important protocol deviations reported, including a description of the protocol deviation. **Listing details example:** 1. Unique Subject ID 2. Subject ID 3. Treatment Group 4. Type of Deviation 5. Deviation Description
Efficacy Endpoints (All Randomized Subjects)	This listing should include all primary and key secondary efficacy endpoints specified in the protocol. **Listing details example:** 1. Unique Subject ID 2. Subject ID 3. Treatment Group 4. Endpoint (Primary and Key Secondary) 5. Survival Duration 6. Randomization Date 7. Event (Y or N) 8. Event or Censoring Date 9. Event or Censoring Date Description
Concomitant Medications (All Treated Subjects)	The listing should include all the prior and concomitant medications collected by the site per protocol. **Listing details example:** 1. Subject ID 2. Treatment Group 3. Start Date 4. Stop Date 5. Medication Name 6. Total Daily Dose 7. Route of Administration 8. Reason

Safety Monitoring (All Treated Subjects)	The listing should include the results of all tests performed for safety monitoring as specified in the protocol. **Listing details example:** 1. Subject ID 2. Treatment Group 3. Test Category 4. Test Name 5. Test Date 6. Visit Number 7. Result 8. Original Unit 9. Normal Range (Low and High)

Table 1. Basic requirements for subject-level data line listing by the site.

CROSS-CHECK SUBJECT-LEVEL DATA LISTING

1. To ensure the quality and accuracy of the listings, generate SAS datasets before generating the RTF files for each site. Validate the generated SAS datasets against corresponding CSR listing datasets to ensure data are in sync with no discrepancies.
2. Listings with special events that did not meet the inclusion criteria, died, or had serious adverse events should be cross-checked against the study report to make sure the number of subjects with special events counts matches. It helps to ensure the accuracy of the reported data.
3. All listings should be converted into PDF and combined by the site for submission. The following items need to be considered when generating and combining the PDF files:
    a. Do not generate a combined PDF in a gigabyte size. Split the PDFs into multiple files if required to make them more accessible and optimized.
    b. Do not split listings for the same site into multiple files. The split should be done by site and not by listings.
    c. Clean combined PDF properties to remove the text from the title, author, subject, and keywords field. Optimize the file for a Fast Web View, and set "Bookmark Panel and Page" as the default layout under the "Initial View" tab.

KNOW YOUR DATA

1. The consented subject listing often has a reason missing for the screen failure subjects. This can be due to the data issue in the DS domain or inclusion/exclusion criteria. This issue should be reported to the data management and clinical team. Identify the possible scenario and explain it in the BIMO reviewer's guide section-4. The reason should be present for all subjects who consented but not randomized or treated.

Site: XXX-XXXX
Listing of All Subjects Consented to Trial
Trial Number: XXX, Site Number: XXXX, Investigator: XXXX, XXX, Country: XXXX

Unique Subject ID	Subject ID	Randomized (Y/N)	Treated (Y/N)	Reason
XXX_XXXXXX	XXXXXX	N	N	← Reason should be present
XXX_XXXXXX	XXXXXX	N	N	

Figure 1. Example of listing without reason for a screen failure subjects.

USUBJID	SUBJID	SITENUM	RANDFL	Reason
XXX_XXXXXX	XXXXXX	N	N	In screening after the database cutoff date of xxxxxxxxx
XXX_XXXXXX	XXXXXX	N	N	In screening after the database cutoff date of xxxxxxxxx

Figure 2. Explanation of screen failure reason in BIMO reviewer's guide (section # 4).

SUMMARY-LEVEL CLINICAL SITE (CLINSITE) DATASET

The summary level clinical site (CLINSITE) dataset is an important part III component of the BIMO package. It should contain all supporting safety and efficacy information for all pivotal studies. The draft guidance for industry and draft specifications to prepare a summary-level clinical site dataset is published by CDER for inspection planning. The dataset is submitted in the form of a transport file (clinsite.xpt), and the data definition (define.xml / define.pdf) is provided to support it.

UNDERSTAND CLINSITE DATASET SPECIFICATION

Understanding the CLINSITE dataset specification requirements and its data source is very important. The clinical site dataset contains 39 variables, which can be categorized further into 4 categories.

The following table represents the variables category, the number of variables under each category, and its data source:

Category	Number of Variables	Name of Variables	Data Source
Study Information	10 (e.g., Study Identifier, Study Title, etc.)	STUDYID, STUDYTL, SPONCNT, SPONNAME, IND, UNDERIND, NDA, BLA, SUPPNUM, ARM	Protocol ADAM.ADSL
Site Information	14 (e.g., Study Site Identifier, Financial Disclosure Amount, etc.)	SITEID, FINLDISC, LASTNAME, FRSTNAME, MINITIAL, PHONE, FAX, EMAIL, COUNTRY, STATE, CITY, POSTAL, STREET, STREET1	Site contact information file provided by the clinical team
Safety Results	8 (e.g., Number of Serious/Non-serious Adverse Events, Number of Deaths, etc.)	SAFPOP, SCREEN, DISCSTUD, DISCTRT, NSAE, SAE, DEATH, PROTVIOL	ADAM.ADSL ADAM.ADAE SDTM.DV, SUPPDV
Efficacy Results	7 (e.g., Treatment Efficacy Results, Treatment Efficacy Results Std Deviation, etc.)	ENDPOINT, ENDPTYPE, TRTEFFR, TRTEFFS, SITEEFFE, SITEEFFS, CENSOR	Protocol ADaM Efficacy Datasets (e.g., ADRS, ADTTE)

Table 2. CLINSITE dataset specification category and data source.

MULTIPLE INVESTIGATORS FOR THE SAME SITE

The site contact information file provides the investigator's name with contact details for each site. It is being used by the CLINSITE dataset to populate the contact details related variables. However, there may be a possibility that a particular site has a multiple investigator history. It is very important to identify the most recent active investigator at the site and provide the information for the same in the dataset. The CLINISTE dataset should contain only one investigator per site, the most recent one. The agency may reject or request a new dataset if more than one investigator's details are provided.

Protocol # - Site #	LASTNAME	FRSTNAME	MINITIAL	PHONE	FAX	EMAIL	COUNTRY	STATE	CITY	POSTAL	STREET
1111-111-1111	Aaaa	Vvvv	A	+1 (999) 999-9999	+1 (999) 999-9999	aaaa.vvvv@123.com	XXX	YYY	ZZZ	12345	XXXX
2222-222-2222	Bbbb	Wwww		+1 (888) 888-8888	+1 (888) 888-8888	bbbb.wwww@123.com	XXX	YYY	ZZZ	12345	XXXX
2222-222-2222	Cccc	Xxxx	B	+1 (777) 777-7777	+1 (777) 777-7777	cccc.xxxx@123.com	XXX	YYY	ZZZ	12345	YYYY
2222-222-2222	Dddd	Yyyy		+1 (666) 666-6666	+1 (666) 666-6666	dddd.yyyy@123.com	XXX	YYY	ZZZ	12345	YYYY
3333-333-3333	Eeee	Zzzz		+1 (555) 555-5555	+1 (555) 555-5555	eeee.zzzz@123.com	XXX	YYY	ZZZ	12345	ZZZZ

Figure 3. Site contact information with multiple investigator contact details for the same site.

In above figure 3, site # 2222 has a history of multiple investigators, and reporting all these details in the CLINSITE dataset is not as per the guideline. In such a scenario, the clinical team needs to reach out to the specific site to confirm the most recent investigator details, which will be displayed in the CLINSITE dataset. For Site # 1111 and 3333, the information is displaying correctly.

CROSS-CHECK THE NUMBERS

It is important to cross-check the counts displayed in the CLINSITE dataset variables against the CSR TLFs or ADaM datasets. Before submission, it helps to identify and resolve the discrepancies between CSR data and the CLINISTE dataset. The following are some of the example checks that programmers should perform to ensure quality deliverables:

- Counts for safety population (SAFPOP) v/s the safety population (TRTFL or SAFFL) counts in the ADSL dataset for each site
- Total number of screened subjects in CLINSITE dataset v/s the "Number of Subjects Screened" in the CSR TLFs
- Counts for non-serious/serious adverse events (NSAE and SAE) v/s the adverse events reported in the ADAE dataset for each site
- Total number of sites reported in the CLINSITE dataset v/s site contact information file received from the clinical team
- Number of deaths reported in the CLINISTE dataset v/s ADSL dataset

EXAMPLE OF COUNT DISCREPANCIES

In the following example, the CLINISTE dataset reports 1 subject in the safety population (Figure 4), while the ADSL dataset frequency gives 2 subjects (Figure 5) in the safety population.

STUDYID	SITEID	ARM	SAFPOP	SCREEN
9999-999	1111	Treatment A	3	5
9999-999	2222	Treatment B	1	2
9999-999	3333	Treatment A	4	6

Figure 4. Safety Population in the CLINSITE dataset.

The FREQ Procedure

SITEID	SAFFL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2222	Y	2	100.00	2	100.00

Figure 5. Safety population in ADSL dataset.

EFFICACY VARIABLES REQUIREMENTS

The CLINSITE dataset should display the following information for each primary and key secondary efficacy endpoints by treatment arm for each site.

- Endpoint (e.g., Progression-free Survival per BICR, Overall Survival)
- Endpoint Type (e.g., Time-to-event, Continuous, Discrete)
- Treatment Efficacy Results
- Treatment Efficacy Result Standard Deviation
- Number of Censored Observations

However, some additional efficacy-related variables with confidence limits may be required based on the study designs. The statistician should identify such requirements, develop the specifications, and provide them to the programmer for development.

Variable	Label	Type	Derivation
TRTEFFC	Trt Efficacy Result Confidence Interval	Char	This variable provides more information regarding treatment result
SITEEFFC	Site-Spec Trt Effect Confidence Interval	Char	This variable provides more information regarding site-specific treatment result

Table 3. Additional efficacy-related variables for a confidence interval.

To ensure quality, programmers must cross-check the number of counts reported for each endpoint and the number of censored subjects against the CSR listing.

STAKEHOLDER COLLABORATION

BIMO deliverables are joint efforts of many stakeholders, as the input data comes from multiple resources. It is very important to open up a good communication channel with all the involved stakeholders to understand the requirements and data collection. The following table identifies the roles and responsibilities of all the stakeholders involved in creating quality BIMO deliverables.

Stakeholders	Roles and Responsibilities
Programmer	• Prepare the environment to create a BIMO package • Gather the requirements to prepare the package • Create a "Subject-level Data Line Listing by Site" outputs • Create a "Site-level Dataset" specification in collaboration with the statistician • Create a CLINSITE dataset • Create a define.xml to support the CLINISTE dataset, and validate it using the P21 • Write a BIMO Reviewer's Guide • Cross-check "Subject-level Data Line Listing by Site" and CLINSITE dataset counts against the ADaM datasets or CSR TLFs
Statistician	• Identifies the protocol-specific "Subject-level Data Line Listing by Site" outputs • Provides the specifications for "Efficacy Results" and "Site-specific Treatment Effect" related variables • Review the CLINISITE dataset and "Subject-level Data Line Listing by Site" outputs
Clinical / Medical Writers	• Prepares clinical-level information • Provides site contact information to prepare the CLINISTE dataset • Confirms the study information (e.g., SPONCNT, SPONNAME, IND/BLA number, etc.)
Data Management	• Resolves the data-related issues in the study • Contact the site to confirm/verify the missing information

Table 4. Roles and responsibilities of all stakeholders involved in BIMO deliverables.

DATA DEFINITION (DEFINE.XML)

The data definition for the CLINSITE dataset is created using the Pinnacle 21 Enterprise tool. The prepared CLINISTE dataset specification should be used to generate the define.xml / define.pdf. The following are some tips for developing a good quality BIMO define a package for the submission.

- Verify the following study-related information
    - Protocol Name
    - Study Name
    - Study Description
    - Standard or Controlled Terminology
    - Supporting Documents (e.g., BIMO Reviewer's Guide)

- Verify the CLINISTE dataset properties
    - Dataset name should be CLINSITE
    - Dataset description, class, structure, and purpose should not be missing

- Dataset variable properties
    - Variable name should not exceed 8 characters
    - The variable label should not exceed 40 characters
    - Significant digits values should be populated when variable type = "float"
    - Variable values should be <= 200. If the values are more than 200, the variables should be split into multiple variables (e.g., DVCAT1, DVCAT2)

- Codelist or controlled Terms
    - ARM codelist values should be matched with codelist values of ADSL.ARM variable
    - Keep only those controlled terms that are used in the CLINSITE dataset

FDA may request the source datasets used to generate the CLINISTE dataset. The examples of the source datasets include, but are not limited to:

- ADSL (Subject-level analysis dataset)
- ADAE (Adverse event-related dataset)
- ADPDEV (Dataset related to protocol deviation)
- ADRS / ADTTE (Efficacy endpoint related datasets)

CONCLUSION

Bioresearch Monitoring (BIMO) submissions are important for the planning of inspections. The technical conformance guide is being used to standardize the agency's BIMO submissions across all therapeutic areas. The pivotal studies require the BIMO package to ensure data integrity and confirm that all clinical trials regulations are met. Although the BIMO package is a standard process and similar across all submissions, the programmer needs to be careful when creating a package. The most up-to-date guidance should be followed and identifies any protocol-specific requirements in collaboration with the stakeholders. This paper identifies how to avoid making common mistakes and provides quick tips on verifying the information to ensure good quality BIMO deliverables.

REFERENCES AND RECOMMENDED READING

- Food and Drug Administration. "Bioresearch Monitoring Technical Conformance guide" Retrieved from https://www.fda.gov/media/85061/download
- Food and Drug Administration. "Standardized Format for Electronic Submission of NDA and BLA Content for the Planning of Bioresearch Monitoring (BIMO) Inspections for CDER Submissions" Retrieved from https://www.fda.gov/files/drugs/published/Standardized-Format-for-Electronic-

Submission-of-NDA-and-BLA-Content-for-the-Planning-of-Bioresearch-Monitoring-%28BIMO%29-Inspections-for-CDER-SubmissionsGuidance-for-Industry.pdf

- Kiran K Kundarapu, Janet C Low, Majdoub Haloui. “Sponsor Considerations for Building a Reviewer’s Guide to Facilitate BIMO Review.” PharmaSUG 2019 - Paper SS-240. Retrieved from https://www.lexjansen.com/pharmasug/2019/SS/PharmaSUG-2019-SS-240.pdf

ACKNOWLEDGMENTS

The authors would like to acknowledge colleagues who provided their feedback and helped us finalize this paper. In particular, we would like to thank Jeff Xia and Amy Gillespie for proofreading and providing critical comments.

CONTACT INFORMATION

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

Name: Chintan Pandya
Enterprise: Merck & Co., Inc.
Address: 126 E. Lincoln Ave
City, State & Zip: Rahway, NJ 07065-4607
Phone: (732) 594-7117
E-mail: chintan.pandya@merck.com
Web: www.merck.com

Name: Himanshu Patel
Enterprise: Merck & Co., Inc.
Address: 126 E. Lincoln Ave
City, State & Zip: Rahway, NJ 07065-4607
Phone: (732) 594-4071
E-mail: himanshu.patel@merck.com
Web: www.merck.com