



Development Process of the Bioresearch Monitoring (BIMO) Package and the Obstacles Successfully Addressed

Murali Neela

Agenda:

- **Biography**
- **Introduction**
- **History of BIMO**
- **BIMO Submission Requirements**
- **Development Process of the BIMO Package**
- **Subject-Level Data Line Listing by Site**
- **Summary-Level Clinical Site Dataset (CLINSITE.XPT)**
- **P21 (DEFINE.XML)**
- **BIMO Data Reviewer's Guide**
- **Challenges**

Biography :

Murali Neela : I am Associate Principal Scientist Stat. Programming at Merck. Over 14 years of experience with specializing in Statistical programming for non-oncology indications. My core strengths include building ADaM datasets, generating TFLs for statistical analysis and reporting, and preparing submission-ready deliverables aligned with CDISC standards and regulatory expectations. I have designed reproducible, validated SAS programs, implemented robust QC processes, and maintained comprehensive metadata and traceability between raw SDTM, ADaM, and outputs.

Sarad Nepal is principal scientist at Merck where he serves a lead programmer for late-stage A&R group 1. In his current role he is responsible for development, planning and management of projects within endocrine, diabetes, CV, and immunology therapy , as well as non-project initiatives, where he applies innovative strategies to enhance processes and increase efficiency.

Madhusudhan Ginnaram is Associate Principal Scientist at Merck, bringing over 13 years of valuable experience in the pharmaceutical industry. He is a lead programmer specializing in late-stage analysis and submission-related deliverables for regulatory agencies globally.

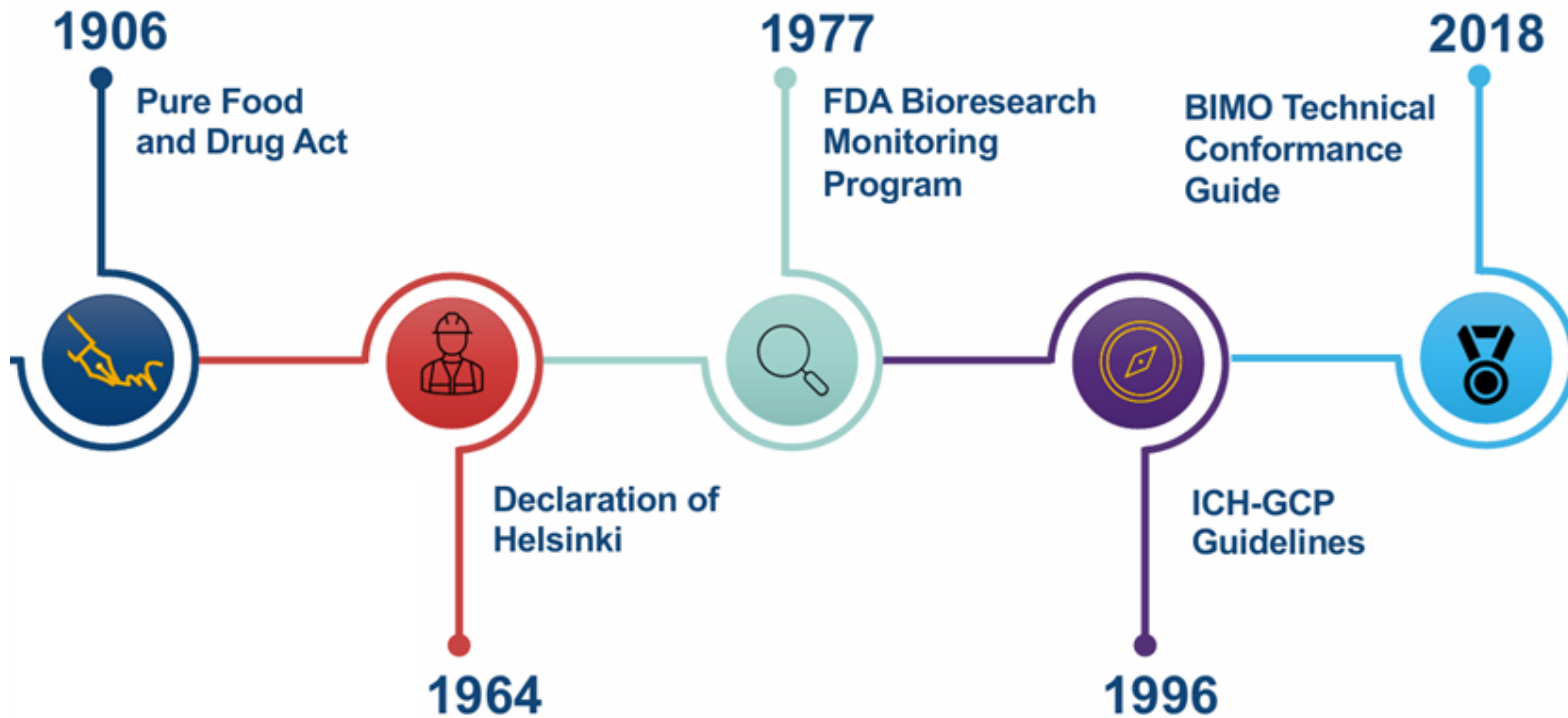
Ramakanth Donthula is a Senior Scientist at Merck with over 13 years of experience in clinical industry. He specializes in SAS programming, ADaM development, statistical graphics, and P21-compliant submission deliverables, contributing high-quality analytical outputs that support regulatory submissions.

Introduction

- **Abstract** :Development Process of the Bioresearch Monitoring (BIMO) Package and the Obstacles Successfully Addressed
- FDA's Bioresearch Monitoring (BIMO) program is a comprehensive program of on-site inspections and data audits designed to monitor all aspects of the conduct and reporting of FDA regulated research. The BIMO Program was established to assure the quality and integrity of data submitted to the agency in support of new product approvals and marketing applications, as well as to provide for protection of the rights and welfare of the thousands of human subjects and animals involved in FDA regulated research. It has become a cornerstone of the FDA preapproval process for new medicines, medical devices, food and color additives, veterinary products and, just recently, new tobacco products introduced to the U.S. consumer.

History of BIMO:

History of BIMO



BIMO Submission Requirements :

Purpose: Planning FDA Bioresearch Monitoring (BIMO) inspections

I. Clinical Study-Level Information

II. Subject-Level Data Line Listings by Clinical Site

- Currently PDF files – Previously known as OSI listings (FDA office of Scientific Investigations)
- Future extracted from SDTM, ADaM Datasets

III. Summary-Level Clinical Site Dataset

- Clinsite.xpt
- Define.pdf

BIMO Data Reviewer's Guide.

Development Process of the BIMO Package:

Inputs needed:

- Protocols, amendments, Annotated CRFs
- ADaM datasets (ADSL, ADAE, efficacy ADaM)
- Site contact information from clinical operations
- Deviation data (DV, SUPPDV)

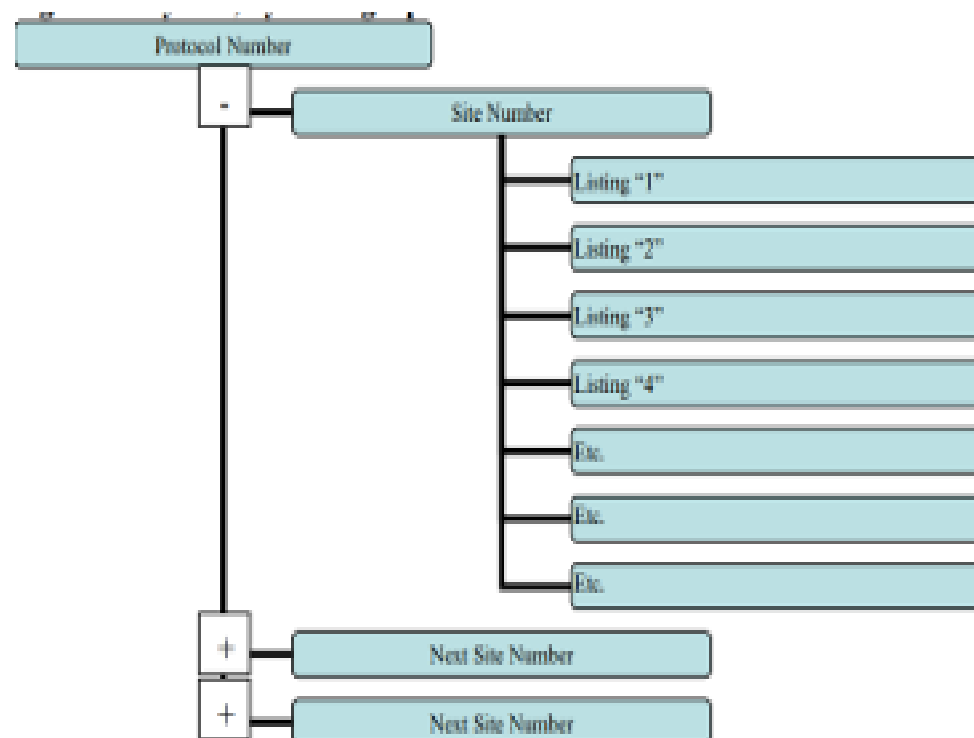
Subject-Level Data Line Listing by Site:

- As part of the NDA/BLA submission package, Bioresearch Monitoring (BIMO) inspection reports will be provided to support inspectional findings and compliance documentation
- The Office of Scientific Investigations (OSI) is responsible for oversight and administration of the Bioresearch Monitoring (BIMO) program.
- Ensure and verify the integrity of efficacy and safety data.
- The Office of Scientific Investigations (OSI)/Bioresearch Monitoring (BIMO) program identifies clinical sites for inspection and assembles inspection packages for inspectors.

Subject-Level Data Line Listing by Site :

BIMO Inspection Package — Components :

1. Consented subjects
2. Treatment assignment
3. Discontinuations
4. Study population
5. Inclusion/exclusion criteria
6. Adverse events
7. Important protocol deviations
8. Efficacy endpoints
9. Concomitant medications
10. Safety monitoring



Subject-Level Data Line Listing by Site:

Site: 0000-000-0100

Listing of All Participants Consented to Trial

Listing of Treatment Assignment (Allocated Participants)

Listing of Reasons Participants Discontinued from Study Medication or Trial (Allocated Participants)

Listing of Study Population (Allocated Participants)

Listing of All Participants Who Did Not Meet Inclusion Criteria or Did Meet Exclusion Criteria

Listing of Adverse Events (Allocated Participants)

Listing of Protocol Deviations (Allocated Participants)

Listing of Primary Efficacy Endpoints (rFAS and rcFAS)

Listing of Prior and Concomitant Medication (Allocated Participants)

Listing of Laboratory Tests (Allocated Participants)

16

16

20

21

22

24

27

28

32

34

45

Site: 0000-000-0100
Listing of Adverse Events
(Allocated Participants)

Trial Number: 0000-000, Site Number: 0100, Investigator: xxx, xxxx, Country: USA

Unique Subject ID	PID	Treatment Group	Onset Date [Relday of Onset]	End Date [Relday of End]	AE MedDRA PT [AE Verbatim]	Intensity [Tox Grade]	Serious	Related	Action Taken	Outcome [Death Date]
No data to report.										
PID= Participant identifier. PID is the allocation number for all randomized participants. For participants who are enrolled but not randomized, PID is the screening number.										
Action Taken: None = DOSE NOT CHANGED, Reduced = DOSE REDUCED, Interrupted = DRUG INTERRUPTED, Discontinued = DRUG WITHDRAWN, Increased = DOSE INCREASED, N/A = NOT APPLICABLE.										
Outcome: Resolved = RECOVERED/RESOLVED, Resolving = RECOVERING/RESOLVING, Sequelae = RECOVERED/RESOLVED WITH SEQUELAE, Not resolved = NOT RECOVERED/NOT RESOLVED.										
Related: Investigator-assessed relationship of the adverse event to study medication.										
Treatment group represents the treatment to which the participant was allocated.										
Relday is the day relative to Visit 2 (enrollment) date.										

Site:0000-000-0900
Listing of Study Population
(Allocated Participants)

Trial Number: 0000-000, Site Number: 0100, Investigator: xxx, xxx, Country: USA

Unique Subject ID	Subject ID	Treatment Group	Completers	Allocated Participants	APaT Participants	FAS Participants	rFAS Participants	rcFAS Participants	FAS Evaluable Participants	PP Participants	Reason for PP Participants Exclusion
0000-000_090000009	90095	XX-0000 XXX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000003	900963	XX-0000 XXX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000004	900242	XX-0000 XXX	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000005	900253	XX-0000 XXX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000006	900274	XX-0000 XXX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000007	900299	XX-0000 XXX	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000008	900359	XX-0000 XXX	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
0000-000_090000009	900368	XX-0000 XXX	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Treatment group represents the treatment to which the participant was actually received. Enrolled participants are equivalent to allocated participants. APaT=All Participant as Treated, FAS=Full Analysis Set, rFAS=Restricted Full Analysis Set, rcFAS=Restricted Contraception Full Analysis Set, PP=Per-Protocol Set.											

Site: 0000-000-0100
Listing of All Participants Who Did Not Meet Inclusion Criteria or Did Meet Exclusion Criteria

Trial Number: 0000-000, Site Number: 0100, Investigator: xxx, xxx, Country: USA

Unique Subject ID	Subject ID	Reason
0000-000_000000000	0010000000	IN07_03: Is able and willing (in the opinion of the investigator) to).

Subject-Level Data Line Listing by Site:

Sponsors generate the 10 required listings using validated SAS macros.

Key requirements:

- Listings must reflect correct populations (screened, randomized, treated)
- All listings should be validated before PDF creation
- Listings must be organized by clinical site, not by study

PDF rules:

- Combine all listings for each site into a single PDF

Do not create excessively large files—split by site ranges if needed

Summary-Level Clinical Site Dataset (CLINSITE.XPT):

The CLINSITE dataset provides one row per site per study and includes:

- Study information (e.g., STUDYID, title, sponsor contact information)
- Site information (e.g., SITEID, investigator name, address, contact details)
- **Safety results:**
Number of screened subjects, treated subjects, protocol violations, deaths, number of serious/non-serious AEs
- **Efficacy results:**
Endpoint, endpoint type, treatment effect, site-specific estimates, censored counts
- Include one investigator per site (the most recent active one)
- Match counts found in ADaM and CSR TLFs
- Adhere to FDA's variable naming, controlled terminology, and length restrictions
- All supporting safety and efficacy information for pivotal studies are included.
- A subject changes sites → explain the convention used
- Endpoints vary across studies → define each endpoint clearly
- The dataset is provided as a transport file (clinsite.xpt). Accompanying metadata and data definitions are included as define.xml and define.pdf to support the dataset.

Summary-Level Clinical Site Dataset (CLINSITE.XPT) :

APPENDIX 3: CLINICAL SITE DATA ELEMENTS SUMMARY LISTING

Table B: Clinical Site Data Elements Summary Listing

Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Notes or Description	Sample Value
1	STUDYID	Study Identifier	Char	String	Study or trial identification number.	ABC-123
2	TITLE	Study Title	Char	String	Title of the study as listed in the clinical study report (limit 200 characters). If the title exceeds 200 characters, provide shortened title and define (e.g., use the abbreviated title from clinicaltrials.gov).	Double blind, randomized, placebo-controlled clinical study on the influence of drug X on indication Y
3	SPONCNT	Sponsor Count	Num	Integer	Total count of sponsors throughout the study. If there was a change in the sponsor while the study was ongoing, with sponsors as defined in § 312.3 (21 CFR 312.3), enter an integer indicating the total count of sponsors. If there was no change in the sponsor while the study was ongoing, enter "1."	1
4	SPONSOR	Sponsor Name	Char	String	Full name of the sponsor organization conducting the study at the time of study completion, as sponsor is defined in § 312.3. If the sponsor name exceeds 200 characters, provide short-form sponsor name and define.	DrugCo, Inc.
5	IND	IND Number	Num	6 digit identifier	IND number. If study not performed under IND, leave blank.	010010
6	UNDERIND	Under IND	Char	String	Value should equal "Y" if study at the site was conducted under an IND (i.e., a Form FDA 1572 was signed by the investigator) and "N" if study was not conducted under an IND at the site (i.e., a Form FDA 1572 was not signed by the investigator).	Y
7	NDA	NDA Number	Num	6 digit identifier	FDA NDA number, if available/applicable. If not applicable, leave blank.	021212
8	BLA	BLA Number	Num	6 digit identifier	FDA identification number for BLA, if available/applicable. If not applicable, leave blank.	123456
9	SUPPNUM	Supplement Number	Num	Integer	Serial number for supplemental application, if applicable. If no information is available, leave blank.	4
10	SITEID	Study Site Identifier	Char	String	Investigator site identifier assigned by the sponsor.	50
11	ARM	Description of Planned Treatment Arm	Char	String	Plain-text label for the name given to an arm or treatment group as referenced in the clinical study report (limit 200 characters). When no arm or treatment group is available due to only screen failure subjects at site, use label "Screen Failure."	Active name and dose (e.g., "Active 25mg"), Comparator product name (e.g., "Drug X"), Placebo, Screen Failure

Variable Index	Variable Name	Variable Label	Type	Controlled Terms or Format	Notes or Description	Sample Value
12	COHORT	Description of Planned Cohort	Char	String	For cohort studies, the plain-text label given to a cohort as referenced in the clinical study report (limit 200 characters). When not a cohort study, leave blank.	A
13	SAFPOP	Number of Subjects in Safety Population	Num	Integer	Total number of subjects in safety population at a given site by treatment arm. When a subject has transferred from one site to another, the applicant should handle reporting of such subjects consistently across sites and include in the define file the reporting convention used. The applicant may opt to further explain the reasons subjects transferred between sites in the BIMO Data Reviewer's Guide, if a guide will be provided.	20
14	EFFPOP	Number of Subjects in Efficacy Population	Num	Integer	Total number of subjects in primary efficacy population as reported in the Clinical Study Report at a given site by treatment arm. Further describe the population reported as EFFPOP (e.g., Per Protocol, Full Analysis Set, Intent to Treat, modified Intent to Treat) in the clinsite define.html. When a subject has transferred from one site to another, the applicant should handle reporting of such subjects consistently across sites and include in the define file the reporting convention used. The applicant may opt to further explain the reasons subjects transferred between sites in the BIMO Data Reviewer's Guide, if a guide will be provided.	18
15	SCREEN	Number of Subjects Screened	Num	Integer	Total number of subjects screened (and consented) at a given site (overall number per site as subjects have not yet been assigned to treatment arm). When a subject has transferred from one site to another, the applicant should handle reporting of such subjects consistently across sites and include the reporting convention used in the define file or the BIMO Data Reviewer's Guide (if provided). The applicant may opt to further explain the reasons subjects transferred between sites in the BIMO Data Reviewer's Guide, if provided.	100
16	DISCSTUD	Number Subjects Discont. Study	Num	Integer	Number of subjects in the safety population who discontinued from the study by treatment arm at a given site.	5
17	DISCRTT	Number Subjects Discont. Study Treatment	Num	Integer	Number of subjects in the safety population who discontinued from the study treatment by treatment arm at a given site.	10
18	ENDPOINT	Primary 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

#	Variable	Type	Len	Label
1	STUDYID	Char	12	Study Identifier
2	TITLE	Char	198	Study Title
3	SPONCNT	Num	8	Sponsor Count
4	SPONSOR	Char	42	Sponsor Name
5	IND	Num	8	IND Number
6	UNDERIND	Char	20	Under IND
7	NDA	Num	8	NDA Number
8	BLA	Num	8	BLA Number
9	SUPPNUM	Num	8	Supplement Number
10	SITEID	Char	41	Study Site Identifier
11	ARM	Char	40	Description of Planned Treatment Arm
12	COHORT	Char	20	Description of Planned Cohort
13	SAFPOP	Num	8	Number of Subjects in Safety Population
14	EFFPOP	Num	8	Number of Subjects in Eff. Population
15	SCREEN	Num	8	Number of Subjects Screened
16	DISCSTUD	Num	8	Number Subjects Discont. Study
17	DISCRTT	Num	8	Number Subjects Discont. Study Treatment
18	ENDPOINT	Char	100	Primary Endpoint
19	ENDPTYPE	Char	10	Primary Endpoint Type
20	TRTEFFR1	Num	8	Treatment Efficacy Result for SAFPOP
21	TRTEFFR2	Num	8	Treatment Efficacy Result for EFFPOP
22	CENSOR1	Num	8	Censored Observations in SAFPOP
23	CENSOR2	Num	8	Censored Observations in EFFPOP
24	NSAE	Num	8	Number of Non-Serious Adverse Events

26	DEATH	Num	8	Number of Deaths
27	IMPDEV	Num	8	Number of Important Protocol Deviations
28	NOIMPDEV	Num	8	Number of Non-Important Prot Deviations
29	FINLDISC	Char	6	Financial Disclosure Amount
30	LASTNAME	Char	100	Investigator Last Name
31	FRSTNAME	Char	100	Investigator First Name
32	MINITIAL	Char	9	Investigator Middle Initial
33	PHONE	Char	22	Investigator Phone Number
34	FAX	Char	16	Investigator Fax Number
35	EMAIL	Char	47	Investigator Email Address
36	COUNTRY	Char	11	Country
37	STATE	Char	40	State
38	CITY	Char	27	City
39	POSTAL	Char	10	Postal Code
40	STREET	Char	96	Street Address

P21 (DEFINE.XML):

The CLINSITE data definition is produced using Pinnacle 21 Enterprise. The finalized CLINSITE dataset specification shall be used to generate the accompanying define.xml and define.pdf metadata files for submission

Datasets

Dataset	Description	Class	Structure	Purpose	Keys	Documentation	Location
CLINSITE	Clinical Site Data Elements	BIMO	One record per site per arm per endpoint		STUDYID, SITEID, ARM, ENDPOINT	All Screened Subjects	clinsite.xpt 

Variable Name	Variable Label	Type	Length	Sig Digits	Format	Codelist / Controlled Terms	Origin	Define Derivation	Core	Core Variable	Multiple Period Only
STUDYID	Study Identifier	Char	12				Assigned		Req		
TITLE	Study Title	Char	198				Assigned		Req		
SPONCNT	Sponsor Count	integer	8				Assigned		Req		
SPONSOR	Sponsor Name	Char	42				Assigned		Req		
IND	IND Number	integer	8				Assigned		Req		
UNDERIND	Under IND	Char	20				Assigned		Req		
NDA	NDA Number	integer	8				Assigned		Req		
BLA	BLA Number	integer	8				Assigned		Req		
SUPPNUM	Supplement Number	integer	8				Assigned		Req		
SITEID	Study Site Identifier	Char	41				Predecessor		Req		
ARM	Description of Planned Treatment Arm	Char	40				Derived		Req		
COHORT	Description of Planned Cohort	Char	20				Assigned		Req		
SAFPOP	Number of Subjects in Safety Population	integer	8				Derived		Req		
EFFPOP	Number of Subjects in Eff. Population	integer	8				Derived		Req		
SCREEN	Number of Subjects Screened	integer	8				Derived		Req		
DISCSTUD	Number Subjects Discont. Study	integer	8				Derived		Req		
DISCTRT	Number Subjects Discont. Study Treatment	integer	8				Derived		Req		

BIMO Data Reviewer's Guide:

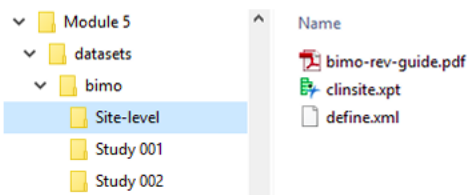
Purpose :

- The BIMO Data Reviewer's Guide provides FDA reviewers and sponsors with standardized expectations and procedures for reviewing clinical trial data submitted to the agency. It helps ensure data integrity, consistency, and regulatory compliance during data review and inspection activities.

Scope:

- The Guide covers data elements, formats, and review workflows relevant to Bioresearch Monitoring (BIMO) activities, including clinical investigator inspections, sponsor/CRO oversight, and data quality assessments across study phases.

Figure 2: Place the Site-Level Dataset **Define** File and BIMO Data Reviewer's Guide in the M5 Folder



Challenges :

File Size and Splitting Issues

Large PDFs slow reviewer navigation.

Split PDFs by site range (e.g., Site **0101**–0200)

Never split listings within a site

Clean PDF metadata (title, author fields)

If the size of the PDF file exceeds **500** megabytes, it should be split into smaller components

Missing or Partial Screen-Failure Data

Screen failure reasons are often incomplete or stored externally.

Source of partial data

How sites tracked screen failures

Limitations of available information

Several listings were missing data for one or more sites.

Insert a blank observation into the dataset.

Append the following code to the PROC REPORT step.

```
compute before / style(lines)=[just=c cell width=8in]; line ' '; line 'No data met the criteria for this listing.'; line ' '; end comp;
```

COUNTRY variable

Previously, TCG used three-letter ISO **3166**-1 alpha-3 country codes. The current TCG uses the GENC (Geopolitical Entities, Names, and Codes) format, aligning with the FDA's adoption of the GENC Standard (U.S. Government profile of ISO **3166**) to ensure conformity with U.S. Government naming and recognition policies.

Challenges :

Special characters in state names:

State names containing special characters (for example, “São Paulo,” “Córdoba,” “Södermanland and Uppland,” “Ôsaka,” and “Île-de-France”) caused errors in the Pinnacle report. These names were provided to us by the Clinical department and were mapped unchanged into the CLINSITE dataset.

Special characters in city names:

City names with special characters (for example, “Córdoba,” “Créteil,” “Ribeirão Preto,” and “São Paulo”) caused errors in the Pinnacle report. These were the original names received from the clinical trial operations team and were mapped unchanged into the CLINSITE dataset.

Financial disclosure amounts:

The symbols in the categories “>= \$25,000” and “< \$25,000” caused issues in the system, despite the underlying amounts being captured correctly.

Postal code issues:

A few non-US postal codes—for example, some from South Korea—caused errors even though they were valid and matched the correct cities.

Dataset	Rule ID	Message	Affected	Changed	Impact	Type	Status	As
CLINSITE	BM0011	Invalid value for FINLDISC variable	324	+4	Medium	Error	Open	
CLINSITE	BM0018	Invalid value in STATE variable	14	0	Low	Error	Open	

Challenges :

Missing or partial Site IDs and site names

Incomplete or incorrect investigator contact details (phone, fax, email, mailing address, postal code)

Investigator roles not specified when multiple staff assigned (Principal vs. Sub-Investigator)

Incomplete or missing site financial data (dollar amounts)

Reminder: Include details only for the Principal Clinical Investigator



Thank you

Name : Murali Neela
Affiliation : Merck & Co., Inc.
Contact Number: (919) 337-7669
E-mail : neela.murali1@merck.com
Website : www.merck.com