

Deriving Phantom Records in BDS Datasets to Support FDA's Standard Safety Figures for Missing Data Analysis

Ruchitbhai Patel, Guowei Wu

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

PHUSE US Connect 2026

Agenda

- Introduction
- Subject Level Data
- Analysis Visit Schedule
- Expected Analysis Visits
- Template ADLB with PHANTOM records
- Analysis
- Conclusion

Introduction

- FDA requires rigorous assessments of observed and missing data in NDAs/BLAs.
 - Primarily affects Laboratory and Vital Sign.
- CDISC ADaM BDS dataset contains recorded measurements.
 - What about Expected but not Collected.
 - Missing expected records necessary to support FDA safety requirements.
- Derive PHANTOM records in BDS dataset to mark scheduled but uncollected assessments.
 - Subject x Treatment x Parameter x Visit
 - PHANTOM records enable traceable and auditable regulatory reporting.

Subject Level Data

Table 1: Subject Level Table with Study Last Day

USUBJID	RANDDT	EOSDT	EOSDY
1	2024-06-19	2024-07-25	36
2	2024-07-03	2024-08-28	56
3	2024-08-23	2024-09-05	13
4	2024-06-15	2024-07-04	19
5	2024-07-06	2024-08-09	34
6	2024-08-12	2024-09-05	24

- EOSDY used as “Study Last Day”.
 - Calculated as difference between EOSDT and RANDDT.
 - Additional factors can be considered for subjects still on study based on study design.
- Ensure clear, documented rules for EOSDY derivation to support reproducibility and regulatory review.

*For presentation purposes, we assume all six subjects completed the study and have EOSDT recorded.

Analysis Visit Scheule

Table 2: Analysis Visit Schedule for Parmeter “PARAM1”.

AVISIT	AVISITN	PARAM	LOWER_BOUND	TARGET_DAY	UPPER_BOUND
Baseline	1	PARAM1	-1	0	1
Visit 2	2	PARAM1	2	22	31
Visit 3	3	PARAM1	32	43	52
Visit 4	4	PARAM1	53	64	73

* The table would be different for each Test/Parameter based on Study Visit Schedule

- For each parameter, guided by study design define visit window.
- Define planned visits (AVISIT, AVISITN) and their windows using LOWER_BOUND, TARGET_DAY, and UPPER_BOUND
- Keep the visit schedule versioned and documented so the mapping and PHANTOM derivations are reproducible for regulatory review.

Expected Analysis Visits

USUBJID	AVISIT	AVISITN	PARAM	LOWER BOUND	UPPER BOUND	RANDDT	EOSDT	EOSDY	Must Have Record
1	Baseline	1	PARAM1	-1	1	2024-06-19	2024-07-25	36	Y
1	Visit 2	2	PARAM1	2	31	2024-06-19	2024-07-25	36	Y
1	Visit 3	3	PARAM1	32	52	2024-06-19	2024-07-25	36	Y
1	Visit 4	4	PARAM1	53	73	2024-06-19	2024-07-25	36	N
2	Baseline	1	PARAM1	-1	1	2024-07-03	2024-08-28	56	Y
2	Visit 2	2	PARAM1	2	31	2024-07-03	2024-08-28	56	Y
2	Visit 3	3	PARAM1	32	52	2024-07-03	2024-08-28	56	Y
2	Visit 4	4	PARAM1	53	73	2024-07-03	2024-08-28	56	Y
3	Baseline	1	PARAM1	-1	1	2024-08-23	2024-09-05	13	Y
3	Visit 2	2	PARAM1	2	31	2024-08-23	2024-09-05	13	Y
3	Visit 3	3	PARAM1	32	52	2024-08-23	2024-09-05	13	N
3	Visit 4	4	PARAM1	53	73	2024-08-23	2024-09-05	13	N
4	Baseline	1	PARAM1	-1	1	2024-06-15	2024-07-04	19	Y
4	Visit 2	2	PARAM1	2	31	2024-06-15	2024-07-04	19	Y
4	Visit 3	3	PARAM1	32	52	2024-06-15	2024-07-04	19	N
4	Visit 4	4	PARAM1	53	73	2024-06-15	2024-07-04	19	N
5	Baseline	1	PARAM1	-1	1	2024-07-06	2024-08-09	34	Y
5	Visit 2	2	PARAM1	2	31	2024-07-06	2024-08-09	34	Y
5	Visit 3	3	PARAM1	32	52	2024-07-06	2024-08-09	34	Y
5	Visit 4	4	PARAM1	53	73	2024-07-06	2024-08-09	34	N
6	Baseline	1	PARAM1	-1	1	2024-08-12	2024-09-05	24	Y
6	Visit 2	2	PARAM1	2	31	2024-08-12	2024-09-05	24	Y
6	Visit 3	3	PARAM1	32	52	2024-08-12	2024-09-05	24	N
6	Visit 4	4	PARAM1	53	73	2024-08-12	2024-09-05	24	N

- Combining **Subject-Level Data** and the **Analysis Visit Schedule**: Identify required records for each subject and parameter.
- "**Must_Have_Record**" is determined using the simple rule $LOWER_BOUND \leq EOSDY \rightarrow Y$ (then flag all preceding visits as Y); otherwise, N (not expected).

Template ADLB with PHANTOM records

Compare Expected Visit with Actual Visits recorded in ADLB. Add PHANTOM records to ADLB as needed.

USUBJID	AVISIT	AVISITN	PARAM	LOWER BOUND	UPPER BOUND	RANDDT	EOSDT	EOSDY	Must Have Record
1	Baseline	1	PARAM1	-1	1	2024-06-19	2024-07-25	36	Y
1	Visit 2	2	PARAM1	2	31	2024-06-19	2024-07-25	36	Y
1	Visit 3	3	PARAM1	32	52	2024-06-19	2024-07-25	36	Y
1	Visit 4	4	PARAM1	53	73	2024-06-19	2024-07-25	36	N



USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
1	ADLB	Baseline	1	Group 1	PARAM1	x.x	
1	ADLB	Visit 2	2	Group 1	PARAM1	x.x	



USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
1	ADLB	Baseline	1	Group 1	PARAM1	x.x	
1	ADLB	Visit 2	2	Group 1	PARAM1	x.x	
1	ADLB	Visit 3	3	Group 1	PARAM1		PHANTOM

USUBJID	AVISIT	AVISITN	PARAM	LOWER BOUND	UPPER BOUND	RANDDT	EOSDT	EOSDY	Must Have Record
2	Baseline	1	PARAM1	-1	1	2024-07-03	2024-08-28	56	Y
2	Visit 2	2	PARAM1	2	31	2024-07-03	2024-08-28	56	Y
2	Visit 3	3	PARAM1	32	52	2024-07-03	2024-08-28	56	Y
2	Visit 4	4	PARAM1	53	73	2024-07-03	2024-08-28	56	Y



USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
2	ADLB	Baseline	1	Group 1	PARAM1	x.x	
2	ADLB	Visit 3	3	Group 1	PARAM1	x.x	
2	ADLB	Visit 4	4	Group 1	PARAM1	x.x	



USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
2	ADLB	Baseline	1	Group 1	PARAM1	x.x	
2	ADLB	Visit 2	2	Group 1	PARAM1		PHANTOM
2	ADLB	Visit 3	3	Group 1	PARAM1	x.x	
2	ADLB	Visit 4	4	Group 1	PARAM1	x.x	

Template ADLB with PHANTOM records

USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
1	ADLB	Baseline	1	Group 1	PARAM1	x.x	
1	ADLB	Visit 2	2	Group 1	PARAM1	x.x	
1	ADLB	Visit 3	3	Group 1	PARAM1		PHANTOM
2	ADLB	Baseline	1	Group 1	PARAM1	x.x	
2	ADLB	Visit 2	2	Group 1	PARAM1		PHANTOM
2	ADLB	Visit 3	3	Group 1	PARAM1	x.x	
2	ADLB	Visit 4	4	Group 1	PARAM1	x.x	
3	ADLB	Baseline	1	Group 1	PARAM1	x.x	
3	ADLB	Visit 2	2	Group 1	PARAM1	x.x	
4	ADLB	Baseline	1	Group 2	PARAM1	x.x	
4	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM
5	ADLB	Baseline	1	Group 2	PARAM1	x.x	
5	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM
5	ADLB	Visit 3	3	Group 2	PARAM1	x.x	
6	ADLB	Baseline	1	Group 2	PARAM1	x.x	
6	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM

New Record
added to ADLB
as needed.

NO Updates
needed for
USUBJID =3.

- USUBJID 1 PHANTOM record for AVISIT = 3 added.
- USUBJID 2 PHANTOM record for AVISIT = 2 added.
- USUBJID 3 No PHANTOM records added.
- USUBJID 4 PHANTOM record for AVISIT = 2 added.
- USUBJID 5 PHANTOM record for AVISIT = 2 added.
- USUBJID 6 PHANTOM record for AVISIT = 2 added.

- Existing records stay as is with available AVAL values.
- For PHANTOM record AVAL is set to Missing.
- Needed variable are add to be ADaM Compliant
- DTYPE= "PHANTOM" aligns with CDISC Controlled Terminology.

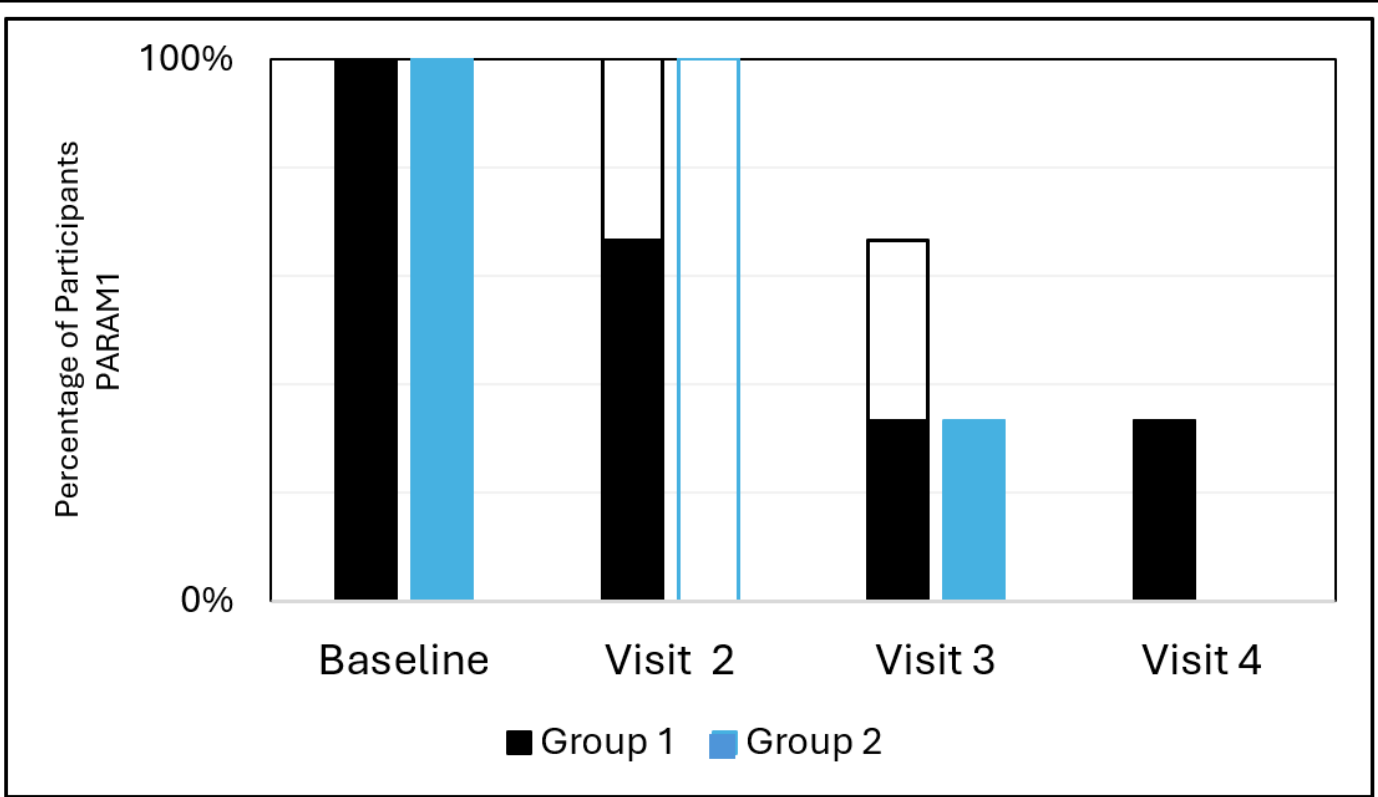
Analysis

USUBJID	DOMAIN	AVISIT	AVISITN	TRT0XP	PARAMCD	AVAL	DTYPE
1	ADLB	Baseline	1	Group 1	PARAM1	x.x	
1	ADLB	Visit 2	2	Group 1	PARAM1	x.x	
1	ADLB	Visit 3	3	Group 1	PARAM1		PHANTOM
2	ADLB	Baseline	1	Group 1	PARAM1	x.x	
2	ADLB	Visit 2	2	Group 1	PARAM1		PHANTOM
2	ADLB	Visit 3	3	Group 1	PARAM1	x.x	
2	ADLB	Visit 4	4	Group 1	PARAM1	x.x	
3	ADLB	Baseline	1	Group 1	PARAM1	x.x	
3	ADLB	Visit 2	2	Group 1	PARAM1	x.x	
4	ADLB	Baseline	1	Group 2	PARAM1	x.x	
4	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM
5	ADLB	Baseline	1	Group 2	PARAM1	x.x	
5	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM
5	ADLB	Visit 3	3	Group 2	PARAM1	x.x	
6	ADLB	Baseline	1	Group 2	PARAM1	x.x	
6	ADLB	Visit 2	2	Group 2	PARAM1		PHANTOM

- Within Each Treatment x Analysis Visit Each Subject is Counted Once.
- Counts in the table below the chart are presented as “remaining in trial with results / remaining in trial (expected)” and reflect inclusion of PHANTOM records in the expected count (denominator) while only observed records contribute to the numerator.

Treatment	Baseline	Visit 2	Visit 3	Visit 4
Group 1	3/3	2/3	1/2	1/1
Group 2	3/3	0/3	1/1	0/0

Analysis



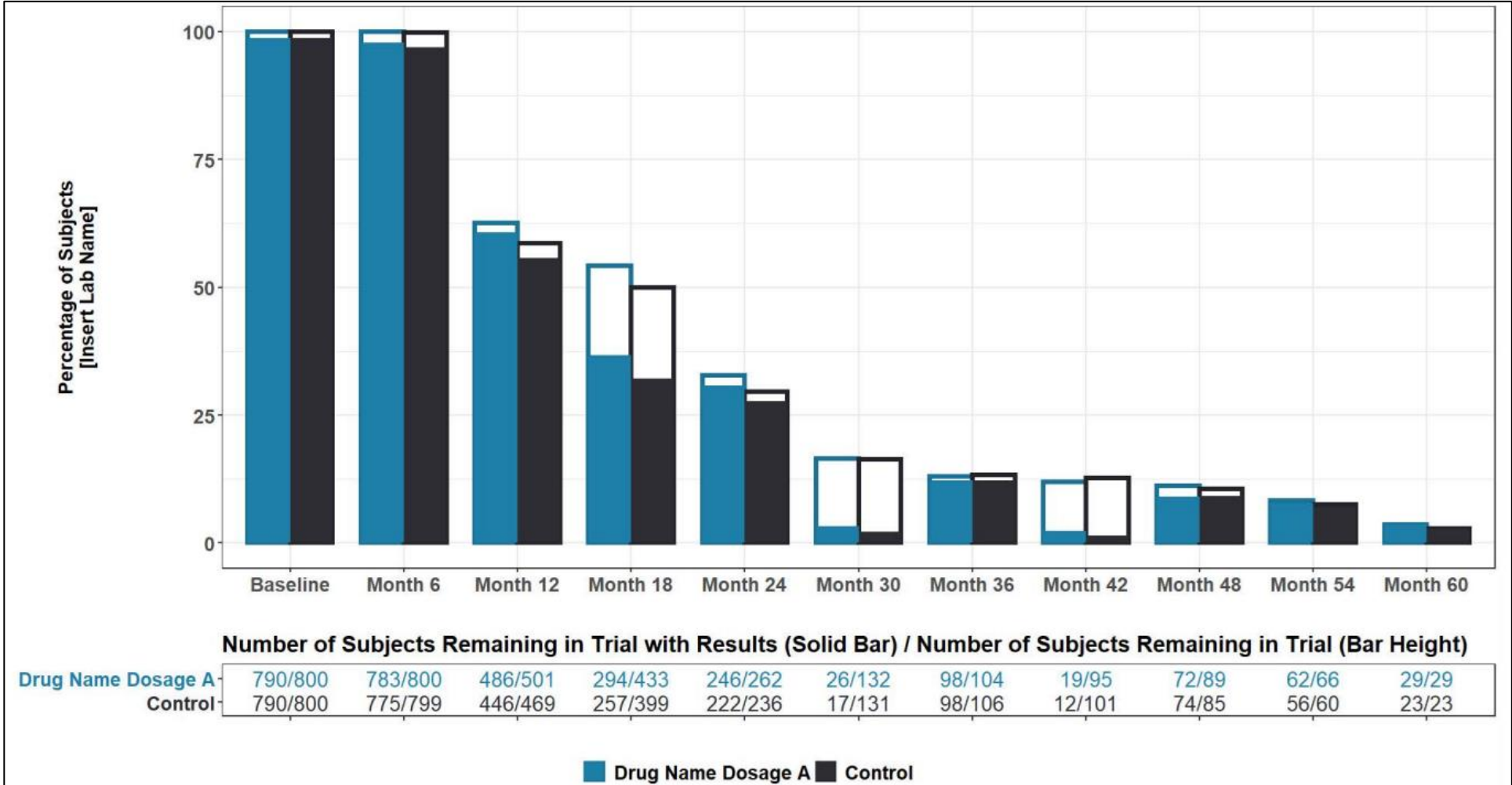
No. of Participants: Remaining in Trial with Results (Solid Bar) / Remaining in Trial (Bar Height)

Treatment	Baseline	Visit 2	Visit 3	Visit 4
Group 1	3/3	2/3	1/2	1/1
Group 2	3/3	0/3	1/1	0/0

- Each bar shows the percentage of participants (within each treatment group) who have an observed (or observed + expected) assessment at the specified visit.
- Percentages are calculated using the number of participants in the analysis population for that treatment group as the denominator. In this example both Group 1 and Group 2 have three participants each.
- Solid Bar: No. of Participants in Trial with Results.
- Bar Height: No. of Participants in Trial

Figure from FDA

Proportion of Subjects Remaining in Trial at Each Visit by Availability of [Insert Lab Value] Results
Safety Pollution



Conclusion

- Explicitly represent PHANTOM rows in BDS (ADLB) to close gap between recorded data and protocol expectations.
- Use subject-level data in combination with visit schedule to determine visits expected for each subject.
- Document steps/practices followed to improve auditability, reduce bias in safety summaries, and strengthen FDA-aligned reporting.
- Future work:
 - Focus on standardizing derivation rules.
 - Automating implementation to ensure broad, consistent adoption.

Questions?

Thank You!

Reference

[standard_safety_tables_and_figures_integrated_guide_2025.pdf](#)

[CDISC ADaM Controlled Terminology](#)