phuse·share
connect·share·advance

Single Day Event

Efficiency in Creating a Good-Quality Submission Package

Anbarasan Durairajan,
Navitas Life sciences
PHUSE Chennai SDE -20SEP2025

Disclaimer

This presentation does not represent views of Navitas Life Sciences and belong solely to the author. I take full responsibility for the work and errors if any.

Agenda

- ❑ Introduction
- ❑ Clinical Study Reports (CSR): Types and Use Cases
- ❑ Key points for efficient SDTM submission
- ❑ Common challenges observed while creating a submission package
- ❑ Key Takeaways
- ❑ Q&A Session

Clinical Study Reports (CSR): Types and Use Cases

CSRs summarize a study's data and outcomes to facilitate the evaluation of a drug's safety and therapeutic effectiveness. Unlike academic journal papers, in which methodological flaws may be glossed over, the CSR provides a detailed description of the study's design and methodology, along with tables, figures, listings, and appendices that further elucidate the data.

.

Type	Description	Use Cases
Full CSR	Complete clinical and statistical description of the study.	Required for regulatory approval (e.g., FDA, EMA), product labeling, marketing authorization
Abbreviated CSR	Condensed version with limited efficacy data but full safety information.	Used for studies not intended to support efficacy claims, such as failed indications or exploratory studies
Synoptic CSR	Expanded synopsis with full safety data, minimal tables or figures.	Appropriate for early-phase safety/tolerance studies not relevant to efficacy evaluation
Supplemental CSR	Adds detail to a full CSR, often referencing the main report.	Used for post hoc analyses, cross-study comparisons, or delayed data inclusion
CSR Amendments	Updates or corrections to a previously submitted CSR.	Used when new data or clarifications are needed after initial submission

Key Points for efficient SDTM submission

aCRF:

- All bookmarks are re-directing to appropriate sections.
- Dual Bookmarks (for VISIT and DOMAIN) should be present for the aCRF

Datasets:

- Check that there is no data with zero observations
- Study description should be aligned across Reviewers Guide-define.xml-TS.xpt and should not be truncated
- Check that there are no datasets >5GB. If datasets exist that are >5GB, then they should be split into smaller datasets

Metadata:

- Ensure alignment and consistency of Meddra / WHO dictionary versions, SDTM IG version and SDTM / MED-RT / SNOMED CT's (and similar) across the equivalent information in the cSDRG, SDTM and define files, Pinnacle21 and Trial Summary (TS) domain.

Dry-run:

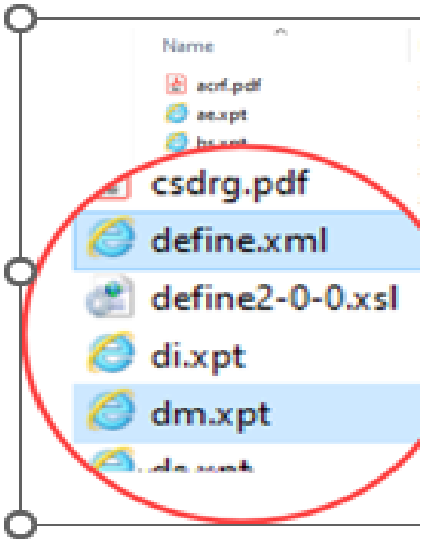
- Do a dry-run on the entire package. Run the complete package through P21 to ensure overall package is eSub compliant.
- Recommend doing it at least one month prior to lock so that you have time to fix any last-minute items if needed prior to the lock

Common challenges observed while creating a submission package

SDTM checks on Technical Rejection criteria:

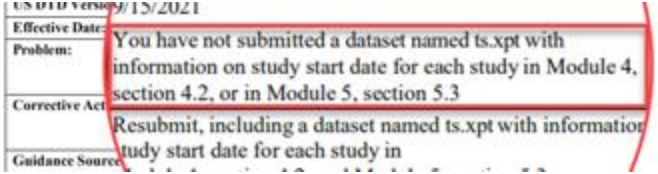
Example1: A DM dataset and define.xml must be submitted

Number:	1736
Group:	General
Description:	DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4. DM dataset, ADSL dataset, define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
Severity Description:	High
US DTD Version	2.01 and 3.3
Effective Date:	TBD
Problem:	You have not submitted Demographic dataset (DM) and define.xml for each study in Module 4, section 4.2. You have not submitted Demographic dataset (DM), Subject level analysis dataset (ADSL), and define.xml for each study in Module 5, section 5.3
Corrective Action:	Resubmit the submission with the Demographic dataset (DM) and define.xml for each study in Module 4, section 4.2. Resubmit including Demographic data, Subject level analysis dataset, and define.xml for each study in Module 5, section 5.3

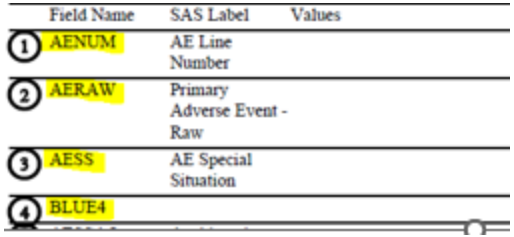


Common challenges observed while creating a submission package

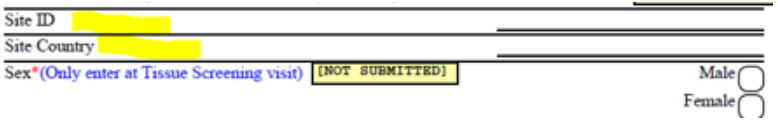
Example 2: SDTM.TS exists with TS.SSTDTC populated with the earliest Informed Consent date in the correct format (yyyy-mm-dd)



Example 3:
Ensure the blank CRF used to do the annotations corresponds to your study data - it should be unique forms only, with no printed field or variable OIDs and no watermarks.

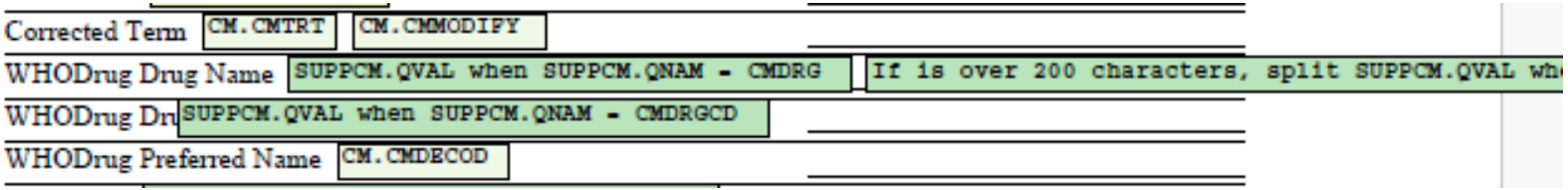


Example 4:
All fields (or pages) should be annotated (if a variable is not submitted annotate with [NOT SUBMITTED])



Common challenges observed while creating a submission package

Example 5: No annotation is truncated, blocking CRF text or off the page



Example 6: Variables referenced in the define.xml should be prefixed by the dataset name

EXSTDY	Study Day of Start of Treatment	integer	Timing	8		Derived Start date (--STDTC) - reference start date (DM.RFSTDTC). Add 1 if start date (--STDTC) is greater than or equal to reference start date (DM.RFSTDTC)

Common challenges observed while creating a submission package

Example 7: Ensure conformance issue explanations documented in the RG's are meaningful and interpretable

Insufficient explanation

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
SD0063	SDTM/dataset variable label mismatch	Warning	PE	1 (6.67%)	Label is fine.

Good – meaningful and informative explanation

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
SD0063	SDTM/dataset variable label mismatch	Warning	PE	1 (6.67%)	Per the SDTMIG, the variable label in the dataset should match the label in the SDTMIG. In SDTMIG v3.2, there are several labels that were erroneously defined to be >40 characters. The variable label for PESTRESC, 'Character Result/Finding in Standard Format', exceeds 40 characters. To avoid truncation, PESTRESC was assigned the label 'Character Result/Finding in Std Format' in the dataset so that it is <= 40 characters due to SAS restrictions.

Common challenges observed while creating a submission package

Example 8:
Ensure alignment and consistency of Meddra / WHO dictionary versions and SDTM IG version across the equivalent information in the cSDRG, SDTM and define files, Pinnacle21 and Trial Summary (TS) domain.

cSDRG

Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
SDTM	•SDTM v1.4 •SDTM-IG v3.2
Controlled Terminology	CDISC SDTM Controlled Terminology, 2018-12-21
Data Definitions	Define-XML v2.0
Medications Dictionary	LOINC 2.69, SNOMED 2020-09-01, UNII 2020-12-14, NDF-RT 2021-01-04
Medical Events Dictionary	MedDRA v23.1

TS

VIEWTABLE: Work.Ts (Trial Summary)

	TSPARMCD	TSPARM	TSVAL	TSVAL1	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
1	ACTSUB	Actual Number of Subjects	114					
2	ADAPT	Adaptive Design	N			C49487	CDISC	2018-11-08
3	ADDON	Added on to Existing Treatments	Y			C49488	CDISC	2018-11-08
4	AEDICT	Adverse Event Dictionary	MedDRA 23.1					
5	AGEMAX	Planned Maximum Age of Subjects			PINF		ISO 21090	
6	AGEMIN	Planned Minimum Age of Subjects	P18Y				ISO 8601	
7	AGESPAN	Age Group	>=18					
8	AGEU	Age Unit	YEARS			C29848	CDISC	2018-11-08
9	CMDICT	WHODrug Dictionary Version	WHODRUG GLOBAL B3 SEPTEMBER 1, 2020					

Key Takeaways

- Data quality is important for efficient submission
- Provide valid explanations for issues in SDRG.
- For blinded trials, ensure unblinding is accurately done.
- An overall P21 validation score of minimum 90% or more should be achieved on the final run.
- Check all tabulation data provided have dataset & variable labels and that they are consistent if referenced within the Define and cSDRG files.
- Check for references of intermediate datasets or external files (*.xls, *.doc) within the RGs and associated define files to ensure such files are documented according to our guidance.

THANK YOU

phuse

The Global Healthcare Data Science Community

Contact Channels
UK +44 1843 609600
office@phuse.global
phuse.global

Social Media
@phusetwitta
/phusebook
/phusetube
/company/phuse