



Vita Global Sciences

FROM RESEARCH TO **REALITY**

Optimizing SDTM and ADaM: Remediation Techniques for Data Accuracy

Kishore Pothuri

What is Remediation?

*The process of identifying, correcting, and improving **data inconsistencies, structural errors, and compliance issues** to ensure that the datasets adhere to latest **CDISC (Clinical Data Interchange Standards Consortium) guidelines** and regulatory requirements.*

Why Remediation Matters?

- Ensures **regulatory compliance** with FDA, PMDA, EMA, and CDISC guidelines.
- Improves **data accuracy** for reliable statistical analysis.
- Reduces risk of **regulatory rejections** due to dataset errors.
- Enhances **regulatory review efficiency** by providing **clean, validated datasets**.

Remediation in SDTM: Purpose & Objectives

- Involves correcting errors in data tabulation, ensuring proper mapping from raw clinical data to SDTM domains, and addressing structural inconsistencies.
- The goal is to ensure that the data meets FDA and CDISC submission requirements.

Key Aspects of SDTM Remediation:

- Correcting **mapping errors** between raw data and SDTM datasets.
- Fixing **structural inconsistencies** such as missing variables or incorrect domain assignments.
- Ensuring compliance with **controlled terminology** (e.g., MedDRA, WHO-DD).

Remediation in SDTM: Purpose & Objectives

- Resolving **cross-domain inconsistencies**, such as mismatches between adverse events (AE) and exposure (EX) datasets.
- Addressing **date format errors** and ensuring proper sequencing of events.

Remediation in ADaM: Purpose & Objectives

- Focuses on refining analysis datasets by ensuring traceability, accuracy in derived variables, and compliance with statistical analysis plans (SAP).
- This step is essential for regulatory submission and reliable statistical analysis.

Key Aspects of ADaM Remediation:

- Ensuring **traceability** from SDTM to ADaM datasets.
- Correcting **derived variable errors** to align with SAP.
- Validating dataset structures (e.g., ADSL should have one record per subject).

Remediation in ADaM: Purpose & Objectives

- Fixing **treatment assignment inconsistencies** across datasets.
- Ensuring **date derivations** are correctly calculated.



Regulatory Agencies: Submission Requirements

- US FDA: Food Drug Administration
- EMA: European Medicines Agency
- Japan PMDA: Pharmaceuticals and Medical Devices Agency
- NMPA CHINA: National Medical Products Administration
- Health Canada
- TGA: Therapeutic Goods Administration of Australia

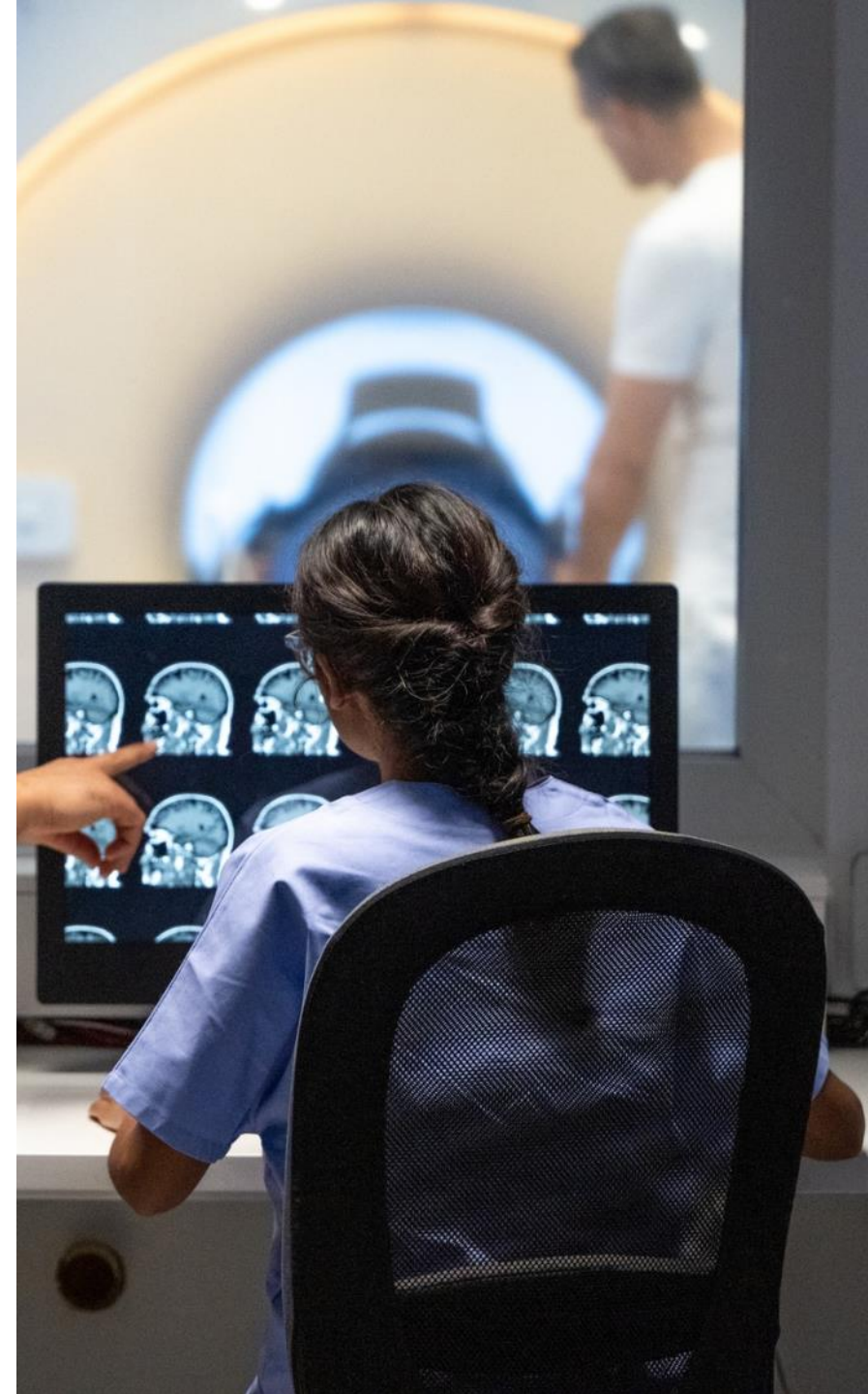
Regulatory Agencies: Submission Requirements

Regulatory Agency	Region	Submission Standard	Submission Type	SDTM Requirements	ADaM Requirements	Submission Format	Key Guidelines
FDA (Food and Drug Administration)	USA	CDISC (SDTM, ADaM)	NDA, BLA, ANDA, IND	Required for all clinical trial data submissions & compliance validated using Pinnacle21	Required for efficacy and statistical analysis & compliance validated using Pinnacle21	eCTD with Define-XML, XPT files	FDA Technical Conformance Guide & 21CFR Part 11 complaint for electronic records
EMA (European Medicines Agency)	Europe	CDISC (SDTM, ADaM recommended)	MAA	Datasets must comply with ICH E3 & E9 guidelines. SDTM Recommended but not mandatory	Datasets must comply with ICH E3 & E9 guidelines. ADaM Recommended but not mandatory	eCTD with Define-XML, XPT files	EMA Guidance on Clinical Data Submission
PMDA (Pharmaceuticals and Medical Devices Agency)	Japan	CDISC (SDTM, ADaM)	NDA, J-NDA	Required for electronic submissions	Required for statistical analysis	eCTD with Define-XML, XPT files	PMDA Technical Conformance Guide
NMPA (National Medical Products Administration)	China	CDISC (SDTM, ADaM recommended)	NDA, IND	Required for NDAs and some INDs	Encouraged but not strictly required	eCTD with Define-XML, XPT files	NMPA Data Submission Guidance
HC (Health Canada)	Canada	CDISC (SDTM, ADaM recommended)	NDS, SNDS, BLA	Highly Recommended but not mandatory	Highly Recommended but not mandatory	eCTD with Define-XML, XPT files	Health Canada Guidance on Data Standards
TGA (Therapeutic Goods Administration)	Australia	CDISC (SDTM, ADaM optional)	NDA, IND	Preferred but not mandatory	Optional	eCTD with Define-XML, XPT files	TGA Clinical Data Submission Guidelines



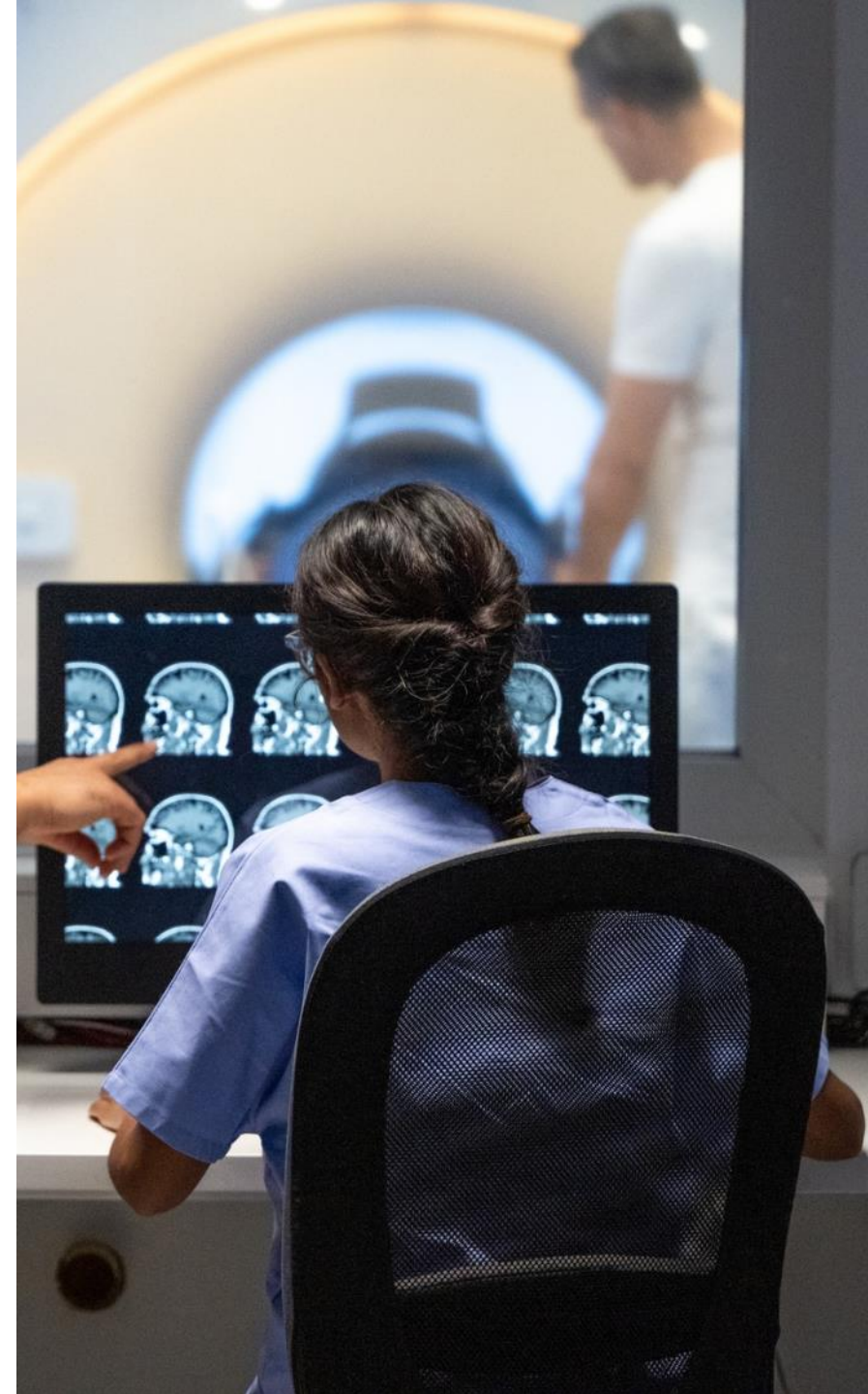
Case Study: Remediation of legacy data in SDTM and ADaM

- Remediation of SDTM (Study Data Tabulation Model) and ADaM (Analysis Data Model) datasets using legacy studies is a critical process in ensuring regulatory compliance and data consistency in clinical trials.
- Legacy studies often predate current CDISC (Clinical Data Interchange Standards Consortium) standards, necessitating the transformation and remediation of historical data to align with modern SDTM and ADaM standards.



Case Study: Remediation of legacy data in SDTM and ADaM

- This process involves mapping legacy data structures to CDISC-compliant formats, addressing any discrepancies, and ensuring that the datasets are complete, accurate, and ready for submission.
- The remediation process enhances the reusability of legacy data, supports integrated analyses across multiple studies, and facilitates more efficient regulatory reviews.



Industry Guidance: SDTM & ADaM Remediation

Is there an industry provided practice or work instructions for SDTM & ADaM remediation?

- ❑ While CDISC and FDA do not provide a direct guide for up-versioning SDTM and ADaM mappings, they provide foundational principles that can be applied. Following these principles ensures alignment with expectations:
- **SDTMIG/ADaMIG Change Log:** Carefully review the documentation for differences between versions.
- **CDISC Community Forums:** Discussions often address challenges and solutions in version migration.
- **Working Groups:** Join CDISC working groups or consult with vendors that specialize in data standards.

Industry Guidance: SDTM & ADaM Remediation

- ❑ Even though there is no gold standard method or practice for remediation of legacy data, there are several industry guidance documents, standards, and tools available for SDTM and ADaM remediation
- ❑ These resources are primarily provided by CDISC (Clinical Data Interchange Standards Consortium) and regulatory agencies such as the FDA, EMA, and PMDA.
- ❑ **Key resources:**
 - **CDISC Guidance Documents:** SDTM Implementation Guide & ADaM Implementation Guide
 - **FDA Guidance:** FDA Technical Rejection Criteria for Study Data & Study Data Technical Conformance Guide
 - **Validation and Remediation Tools:** Pinnacle 21 Enterprise & Pinnacle 21 Community, SAS Programming

Approach Followed for Remediation of Legacy Data:

❑ Objectives:

- Conversion of legacy data from SDTM IG version 3.1.3 to 3.4
- Use the latest SDTM Controlled Terminology
- Generation and Validation of SDTM define.xml v 2.1.0
- Conversion of legacy data from ADaM IG version 1.1 to 1.3
- Use the latest ADaM Controlled Terminology
- Generation and Validation of ADaM define.xml v 2.1.0
- Generation of other regulatory submission components such as SDTM aCRF, cSDRG, ADRG and any other supporting documentation.

Approach Followed for Remediation of Legacy Data:

❑ Steps in Legacy Data Remediation

➤ Gap Analysis:

- Compared legacy datasets against current SDTM/ADaM standards

➤ Mapping and Transformation:

- Updated variable names, formats, controlled terminology, and domain structures.

➤ Validation:

- Used tools like Pinnacle 21 to validate datasets and ensure compliance.

➤ Documentation Update:

- Revised Define-XML, annotated CRFs, and reviewer guides.

Issue Summary				
Source	Pinnacle 21 ID	Message	Severity	Found
GLOBAL				
	SD1111	Missing SE dataset		1
AE				
	SD0057	SDTM Expected variable AEBDSYCD not found		1
	SD0057	SDTM Expected variable AEHLGT not found		1
	SD0057	SDTM Expected variable AEHLGTCD not found		1
	SD0057	SDTM Expected variable AEHLT not found		1
	SD0057	SDTM Expected variable AEHLTCD not found		1
	SD0057	SDTM Expected variable AELLT not found		1
	SD0057	SDTM Expected variable AELLTCD not found		1
	SD0057	SDTM Expected variable AEPTCD not found		1
	SD0057	SDTM Expected variable AESOC not found		1
	SD0057	SDTM Expected variable AESOCCD not found		1
	SD1077	Regulatory Expected variable EPOCH not found		1
DM				
	SD0057	SDTM Expected variable ACTARM not found		1
	SD0057	SDTM Expected variable ACTARMCD not found		1
	SD0057	SDTM Expected variable ACTARMUD not found		1
	SD0057	SDTM Expected variable ARMNRS not found		1
	SD0057	SDTM Expected variable DTHDTC not found		1
	SD0057	SDTM Expected variable DTHFL not found		1
	SD0057	SDTM Expected variable RFICDTC not found		1
	SD0057	SDTM Expected variable RFPENDTC not found		1
	SD0057	SDTM Expected variable RFXENDTC not found		1
	SD0057	SDTM Expected variable RFXSTDTC not found		1
	SD1082	Variable length is too long for actual data		10
	SD1083	Missing DMDY variable when DMDTC variable is present		1

➤ **P21Report:** Legacy SDTM data in SDTM IG

version 3.1.2 vs SDTM IG version 3.4

➤ **Key Findings :** Few examples

- Subject Elements (SE) domain missing
- Adverse Events (AE) missing the coding variables such as AEBDSYCD, AEHLT etc.
- Regulatory Expected Variable EPOCH is not found
- Demographics (DM) missing expected variables such as ACTARM/CD, ACTARMUD, ARMNRS, RFICDTC, RFPENDTC etc.

Issue Summary				
Source	Pinnacle 21 ID	Message	Severity	Found
ADAE				
	AD0018	Variable label mismatch between dataset and ADaM standard		2
	AD0047	Required variable is not present		14
ADSL				
	AD0019	COMPLFL subject-population flag value is null		21
	AD0320	Non-standard dataset label		1
	SD0059	Define.xml/dataset variable type mismatch		13
	SD0060	Variable in dataset is not present in define.xml		2
	SD1082	Variable length is too long for actual data		6
	SD1324	Define.xml/dataset variable label mismatch		1

- **P21Report:** Legacy ADaM data in ADaM IG version 1.1 vs ADaM IG version 1.3
- **Key Findings :** Few examples
 - Analysis Dataset of Adverse Events (ADAE) missing the coding variables such as AEBDSYCD, AEHLT etc.
 - Analysis Data Subject Level (ADSL) has the COMPLFL value as null
 - Analysis Data Subject Level (ADSL) has variables which are not present in define.xml

Remediation Plan

- Based on the GAP analyses performed, a combined remediation plan document for SDTM & ADaM was developed to document all the findings, planning, strategies, objectives, scope and resolutions etc. for remediation.
- This can be a living document but needs to be updated simultaneously along with remediation updates.
- After many iterations, the final version was agreed by all the stake holders and finally signed off for documentation.

Remediation Plan

➤ **Contents of a Remediation Plan can include the below but not limited to:**

- Introduction
- Objectives
- Scope
- Current State Assessment: SDTM & ADaM
- Stakeholder Input
- Comparison with Best Practices
- Documentation of Findings

Remediation Plan

➤ **Contents of a Remediation Plan can include the below but not limited to:**

- Remediation Strategy: aCRF, SDTM & ADaM
- Action Plan
- Resources Required
- Timelines
- Communication Plan
- Potential Constraints/Risks
- Quality Control

Remediation Plan

➤ **Contents of a Remediation Plan can include the below but not limited to:**

- Documentation
- Approval
- Appendix

Challenges

- Structural Differences Between Legacy Data and SDTM/ADaM Standards
- Incomplete or Missing Data Elements
- Controlled Terminology Updates
- Traceability Issues Between Raw, SDTM, and ADaM Datasets
- Handling Unstructured or Free-Text Data
- Regulatory Compliance and Validation Challenges
- Study Design Inconsistencies
- Define-XML and Metadata Challenges
- Lack of Documentation and Institutional Knowledge

Conclusion

- Remediating legacy data into **CDISC-compliant SDTM and ADaM formats** is challenging but crucial for regulatory submissions and integrated analyses.
- The key to success is a **structured approach**, leveraging **automation, expert validation, and adherence to the latest CDISC guidance**.

Questions?

Contact Information

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

Author Name: Kishore Pothuri

Company: Vita Global Sciences

Address: 281 Winter St, First Floor

City / Postcode: Waltham, MA, 02451

Work Phone: 781-833-0259

Email: kpothuri@softworldinc.com

Web: www.vitaglobalsciences.com