Optimizing Vendor Data Mapping: A DTS-Driven Approach to SDTM Readiness

Chandan Patel Malyala, Prasad Attini, Cytel Inc., Hyderabad, India

Problem and Context

External vendor data (e.g., Labs, ECG, PK) are sourced from multiple providers with changing structures and conventions. Despite the availability of Data Transfer Specifications (DTS), external/vendor datasets frequently show inconsistencies in variable presence, attributes, visit labels, and codelists. Late detection of structural issues results in repeated data transfers, manual validation, increasing rework, and extended timelines.

This poster summarizes the need for an early, and metadata-driven validation approach that supports external data compliance with approved DTS prior to downstream SDTM mapping and readiness.

Requirements and Challenges

- Apply consistent validation across all external data domains
- Reduced dependency on hard-coded checks
- Use approved DTS as the authoritative metadata source
- Early and transparent issue detection
- Manage various vendor formats and changing DTS versions

Technical Solution (DTS-driven Macro)

The macro runs one external file at a time and generates a consolidated report by external data type.

A consolidated report is generated and shared with DM/Vendor for resolution

External Data → DTS Metadata tables → DTS driven validation Macro → Issues Report and Resolution → SDTM Mapping

Based on approved DTS received, create 3 excel sheets manually:

- DTS variable metadata
- DTS Codelists
- DTS visit mapping

- Attribute Compliance (Length / Label / Type)
- Variables added / dropped
- Codelist consistency
- Visit mapping validity
- Data Quality (Non-printable characters, Duplicate records)
- Comparison of two external files (xlsx/csv/sas7bdat)
  - Records added / removed
  - Value changes for same key variables

Results

➢ Check for missing or unexpected variables in the external dataset

➢ Verify variable attributes (type, length, label) against the Data Transfer Specification

Domain	Issue Category	Variabe Name	Label - DTS	Type - DTS	Length - DTS	Label - External	Type - External	Length - External	Issue
QS	Variable existence & attribute checks	DOMAIN	Domain Abbreviation	Char	2	Domain Abbreviation_	Char	2	LABEL MISMATCH
QS	Variable existence & attribute checks	QSCAT	Category of Question	Char	25	Category of Question	Char	16	LENGTH MISMATCH
QS	Variable existence & attribute checks	SITE	Site Identifier	Char	6				VARIABLE MISSING IN EXTERNAL
QS	Variable existence & attribute checks	QSDESCR				Description	Char	100	NEW VARIABLE IN EXTERNAL
QS	Variable existence & attribute checks	QSDESCR				Description	Char	100	TYPE MISMATCH

➢ Validate VISIT → VISITNUM mapping according to defined rules

➢ Confirm VISITNUM logical consistency across records

Domain	Issue Category	Visit - DTS	Visit - External	Visitnum - DTS	Visitnum - External	Issue
QS	Visit mapping consistency		Visit 4		5	VISIT NOT IN DTS MAP
QS	Visit mapping consistency	Visit 3		4		VISIT IN DTS MAP BUT NOT IN EXTERNAL
QS	Visit mapping consistency	Unscheduled	Unscheduled	8	6	VISITNUM INCONSISTENT WITH DTS

➢ Identify duplicate records within the external data

Domain	Issue Category	QSTEST	VISIT	QSDTC	Number of Duplicates	Issue
QS	Data Quality	Total Score	Visit 2	30JUN23:10:32:10	2	DUPLICATE RECORDS
QS	Data Quality	Total Score	Visit 4	02JUL23:12:10:25	3	DUPLICATE RECORDS

➢ Ensure no non-printable characters are present

Domain	Issue Category	Variable name	Value	Issue
QS	Data Quality	QSTEST	Mean Score for S® Sleep	NON-PRINTABLE CHARACTER
QS	Data Quality	QSDESCR	Where10 indicates High and 0® indicates Low	NON-PRINTABLE CHARACTER

➢ Codelist consistency

Domain	Issue Category	Test name - DTS	Test name - External	Test code - DTS	Test code - External	Issue
QS	Test name/code consistency		QS Test name 8			TEST NOT IN DTS MAP
QS	Test name/code consistency	QS Test name 5				TEST NOT IN EXTERNAL
QS	Test name/code consistency				QSTEST8	TESTCD NOT IN DTS MAP
QS	Test name/code consistency			QSTEST5		TESTCD NOT IN EXTERNAL

➢ Compare current vs. previous external files to detect structural or content changes

In Previous version

Domain	Issue Category	QSSEQ	QSTESTCD	QSORRES	VISITNUM	VISIT	QSDTC
QS	Version Comparisi	4	TSCORE	75	1	Baseline	10OCT24:10:30:25
QS	Version Comparisi	23	TSCORE	47	2	Visit 1	11OCT24:12:20:15
QS	Version Comparisi	42	TSCORE	88	6	Unscheduled	12OCT24:10:15:10

In New/Current version

Domain	Issue Category	QSSEQ	QSTESTCD	QSORRES	VISITNUM	VISIT	QSDTC
QS	Version Comparisi	4	TSCORE	75	1	Baseline	10OCT24:10:30:25
QS	Version Comparisi	23	TSCORE	88	2	Visit 1	12OCT24:10:15:10
QS	Version Comparisi	42	TSCORE	85	3	Completion	14OCT24:09:20:09
QS	Version Comparisi	61	TSCORE	47	6	Unscheduled	11OCT24:12:20:15

Difference in New vs Previous version file comparison

Domain	Issue Category	QSSEQ	QSTESTCD	QSORRES	VISITNUM	VISIT	QSDTC	Issue
QS	Version Comparisi	23	TSCORE	47	2	Visit 1	11OCT24:12:20:15	DELETED/UPDATED RECORD COMPARED TO PREVIOUS TRANSFER
QS	Version Comparisi	23	TSCORE	88	2	Visit 1	12OCT24:10:15:10	NEW RECORD ADDED IN NEW TRANSFER
QS	Version Comparisi	42	TSCORE	85	3	Completion	14OCT24:09:20:09	NEW RECORD ADDED IN NEW TRANSFER
QS	Version Comparisi	61	TSCORE	47	6	Unscheduled	11OCT24:12:20:15	NEW RECORD ADDED IN NEW TRANSFER
QS	Version Comparisi	42	TSCORE	88	6	Unscheduled	12OCT24:10:15:10	DELETED/UPDATED RECORD COMPARED TO PREVIOUS TRANSFER

Lessons Learned (Benefits and Limitations)

Benefits

- Improved Data Quality: Early detection of inconsistencies ensures reliable and accurate data.
- Operational Efficiency: Significant reduction in manual effort and faster turnaround.
- Adaptability: Can be applied across multiple studies and adapted to study specific DTS requirements.
- Future-Ready: Supports SDTM readiness for downstream activities.
- Compliance: Adherence to DTS from the start.

Limitations

- Effectiveness depends on DTS quality.
- Metadata compliance does not replace reconciliation against CRF data.
- Initial transfer compares External data (metadata) only
- Subsequent transfers compare External data metadata and previous/new files at value level

ABBREVIATIONS

DTS = Data Transfer Specifications

Contact Details

Chandan Patel Malyala: chandanpatel.malyala@cytel.com, Prasad Attini: prasad.attini@cytel.com