

Improving Traceability for Complex Algorithms in ADaM Datasets

Priya Saradha, Levstat® LLC

ABSTRACT Currently, ADaM datasets implement traceability using the variable triplet SRCSEQ, SRCDOM, and SRCVAR to the maximum extent possible. These variables enable the user to identify source records from SDTM or other ADaM datasets. The usage of these variables has been largely limited to identifying a single record from the source dataset. Traceability becomes a challenge when multiple records from a single dataset or multiple datasets contribute to the derivation of a single record in the output dataset. To address this challenge, we adapted the usage of SRCSEQ by creating it as a character variable, SRCSEQS, to include all the sequence numbers and their respective sources contributing to a specific record. This paper intends to discuss this simple, yet effective, approach in detail providing implementation examples.

CURRENT CHALLENGE Due to the increasing complexity of data analysis in clinical trials, there is a huge demand to bring information from multiple SDTM and/or ADaM datasets which poses a great challenge to implement traceability. There are two common situations where traceability is a challenge – Multiple records from any single data or Multiple datasets contributing to the derivation of any single analysis result/value. Examples - Cumulative dose computation in ADEX, Imputation of the AVAL in ADQS, Creation of total score based on a specific algorithm in ADQS, Calculation of time-to-event variable in ADTTE.

LIMITATION SRCSEQ is defined as a numeric variable limiting it to store one value.

SOLUTION We define **SRCSEQS** as a character variable, enabling it to accept more than one value. SRCSEQS is defined as a combination of individual blocks separated by a comma (,).

<source>-<sequence number/range>
[, <source>-<sequence number/range>]

where the sequence number is an individual value and the range is bound by a lower and upper value separated by '-'.
 Examples -

a. CM-29, CM-40

Here CM records are identified by CMSEQ values 29 and 40 contribute to the analysis result/value of the current record in the analysis dataset.

b. TR-11-13, RS-32

The range 11-13 from TR indicates that records with TRSEQ as 11, 12, and 13 are used. Similarly, the record from RS with RSSEQ=32 is used to derive the current analysis value along with the records from TR domain.

ADEX (EXPOSURE ANALYSIS DATASET) - This example illustrates computation of average dose of study drug within the study for each patient. The below snapshot from Exposure domain, lists the dose administered for each patient/study drug combination.

STUDYID	USUBJID	EXSEQ	EXTRT	EXDOSE	EXDOSU	EXSTDTC	EXENDTC
XYZ	XYZ-01-001	10	Drug X	5	mg	2018-04-17	2018-04-28
XYZ	XYZ-01-001	11	Drug X	10	mg	2018-04-29	2018-05-03
XYZ	XYZ-01-001	14	Drug Y	5	mg	2018-05-04	2018-05-10
XYZ	XYZ-01-001	16	Drug Y	15	mg	2018-05-11	2018-05-20

SRCSEQS Usage

STUDYID	USUBJID	PARAMCD	PARAM	AVAL	DTYPE	SRCSEQS
XYZ	XYZ-01-001	AVGDOSX	Average dose of Drug X (mg)	7.5	AVERAGE	EX-10-11
XYZ	XYZ-01-001	AVGDOSY	Average dose of Drug Y (mg)	10	AVERAGE	EX-14,EX-16

ADQS (QUESTIONNAIRE ANALYSIS DATASET) - This example illustrates how SRCSEQS can be used to reference derived records within the same ADaM dataset. The below records are part of QS domain where 4 Scores are collected for each visit/patient combination.

STUDYID	USUBJID	QSSEQ	QSTESTCD	QSTEST	QSSTRESN	QSSTRESC	VISIT
XYZ	XYZ-01-001	20	SC01	Score 1	4	4	VISIT 5
XYZ	XYZ-01-001	21	SC02	Score 2	5	5	VISIT 5
XYZ	XYZ-01-001	22	SC03	Score 3	.	.	VISIT 5
XYZ	XYZ-01-001	23	SC04	Score 4	1	1	VISIT 5

SRCSEQS Usage

USUBJID	ASEQ	PARAMCD	PARAM	AVISIT	AVAL	DTYPE	SRCDOM	SRCVAR	SRCSEQS
XYZ-01-001	32	SC01	Score 1	VISIT 5	4	.	QS	QSSTRESN	QS-20
XYZ-01-001	33	SC02	Score 2	VISIT 5	5	.	QS	QSSTRESN	QS-21
XYZ-01-001	34	SC03	Score 3	VISIT 5	.	.	QS	QSSTRESN	QS-22
XYZ-01-001	35	SC03	Score 3	VISIT 5	5	LOCF	QS	QSSTRESN	QS-21
XYZ-01-001	36	SC04	Score 4	VISIT 5	1	.	QS	QSSTRESN	QS-23
XYZ-01-001	37	TSCORE	Total Score	VISIT 5	15	.	ADQS	AVAL	ADQS-32-33, ADQS-34-36

CONCLUSION This approach increases traceability for complex situations and makes the ADaM leaner for submission purposes.