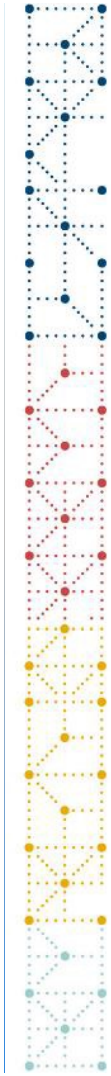
WITH STANDARDS – UNLOCK THE POWER OF DATA

cdisc

2022
US
INTERCHANGE
26-27 OCTOBER | AUSTIN

Demonstrating Traceability in ADaM Datasets

Presented by Wayne Zhong, Consultant, Accretion Softworks LLC

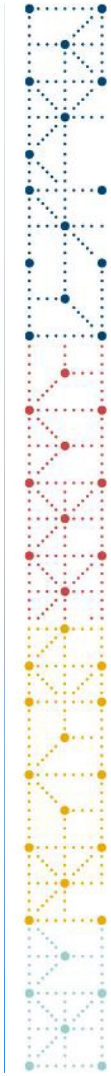


Analysis Data Model (ADaM) Examples of Traceability

Version 1.0, 2022-05-12

Published!

cdisc



Why Do We Need Traceability?

cdisc

Why Do We Need Traceability?

- Clinical studies are conducted to demonstrate new drugs and therapies are safe and effective
- Study data and safety/efficacy results are submitted to agencies
- Traceability is the documentation of steps taken between collected data and the analysis results
- Traceability ensures the analysis results are verifiable

Why Do We Need Traceability? Types of Traceability

- Sample efficacy analysis:

Table xx.x Primary Efficacy Endpoint

ITT Population

	Drug n (%) (N=8000)	Control n (%) (N=8000)	Odds Ratio	P-Value
Occurrence of Primary Study Disease at 2 Years	8 (0.1%)	64 (0.8%)	0.1241	< 0.0001

Why Do We Need Traceability?
Types of Traceability

- Analysis Result Metadata, Metadata Traceability

Display	Table xx.x Primary Efficacy Endpoint
Analysis Result	Occurrence of Primary Study Disease at 2 Years
Analysis Parameter(s)	PARAMCD = "PRI" (Primary Efficacy Endpoint)
Analysis Variable(s)	AVAL (Analysis Value)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADEF [PARAMCD = "PRI" and ITTFL = "Y"]
Documentation	SAP Section 4.1
Programming Statements	[SAS Version 9.2] proc freq data=adef(where=(ittfl='Y' and paramcd='PRI')); table trt01pn*aval; exact or; run;

Why Do We Need Traceability? Types of Traceability

- ADEF Dataset Records, Data Point Traceability

USUBJID	SRCDOM	SRCSEQ	PARAMCD	PARAM	AVAL	AVALC
XYZ-01-001	PF	2	PRI	Primary Efficacy Endpoint	0	DISEASE
XYZ-01-002	LB	52	PRI	Primary Efficacy Endpoint	0	DISEASE
XYZ-01-003			PRI	Primary Efficacy Endpoint	1	NO DISEASE
XYZ-01-004			PRI	Primary Efficacy Endpoint	1	NO DISEASE

Why Do We Need Traceability?
Types of Traceability

- ADEF Variable Metadata, Metadata Traceability

Name	Variable Label	Variable Metadata
USUBJID	Unique Subject Identifier	ADSL.USUBJID
SRCDOM	Source Data	If AVAL=0, identify whether the corresponding record is from PF or LB SDTM domain
SRCSEQ	Source Sequence	If AVAL=0, copy over the corresponding PFSEQ or LBSEQ value from the corresponding record
PARAMCD	Parameter Code	Set to "PRI"
PARAM	Parameter	Set to " Primary Efficacy Endpoint"
AVAL	Analysis Value	If subject has a biopsy record in PF where PFTEST="BIOMARKER 1" and PFSTRESC="PRESENT" then set AVAL=0. Else if subject does not have any biopsy records in PF and has an enzyme record in LB where LBTEST="ENZYME A" and LBSTRESC="POSITIVE" then set AVAL=0. (note: if a biopsy absent record is present, do not check enzyme test records) Otherwise set AVAL=1 Refer to SAP section 4.1 for more details
AVALC	Analysis Value (C)	If AVAL=0 then set AVALC="DISEASE" If AVAL=1 then set AVALC="NO DISEASE"

Why Do We Need Traceability?

- Traceability can answer questions
  - How is a result computed?
  - Where is the data supporting a result?
  - How is the data derived?
  - Which data points in ADaM and SDTM support each subject?
- Without traceability, a reviewer must
  - Examine submitted program code for answers
  - Request meetings with the sponsor for clarification

Why Do We Need Traceability?

- This presentation provides four examples from the ADaM Traceability Examples document

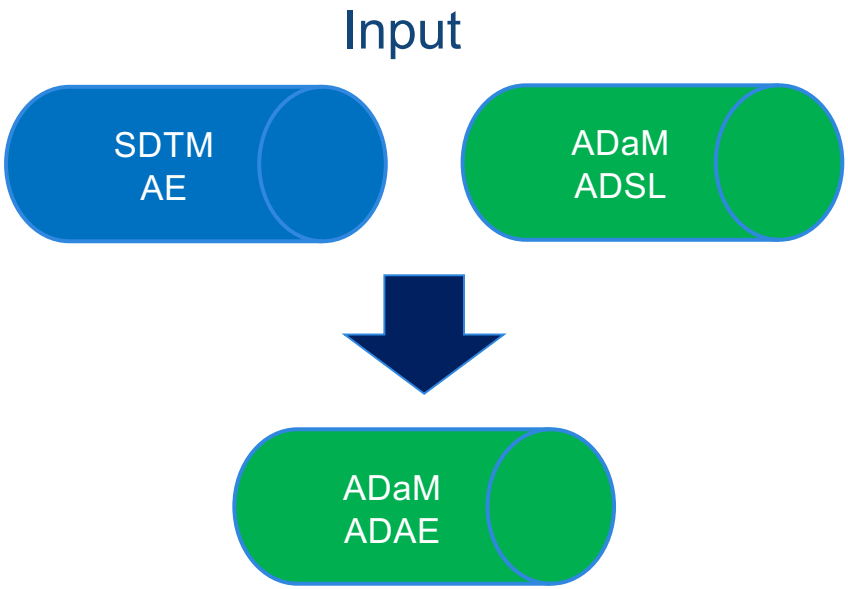
- Please note all examples (including data structures, algorithms, data flows, table shells) are for illustration purposes and are not meant to represent a standard way of analyzing data

Example 1: General OCCDS Traceability

General OCCDS Traceability Analysis Need

- Occurrence analysis is the counting of subjects with a given event, and often includes dictionary coding categories
- Typical analysis includes Adverse Events, Concomitant Medications, and Medical History
- The structure for occurrence analysis dataset is usually one record per record in the corresponding SDTM domain

General OCCDS Traceability Data Flow



General OCCDS Traceability Metadata

- Dataset Metadata for ADAE

Dataset Name	Dataset Description	Dataset Structure	Class of Dataset
ADAE	Adverse Events Analysis Dataset	One record per record in SDTM domain AE (USUBJID AETERM ASTDT AENDT AESEQ).	OCCURRENCE DATA STRUCTURE

- Variable Metadata for ADAE

Name	Variable Label	Variable Metadata
STUDYID	Study Identifier	AE.STUDYID
USUBJID	Unique Subject Identifier	AE.USUBJID
AESEQ	Sequence Number	AE.AESEQ
AETERM	Reported Term for the Adverse Event	AE.AETERM
AEDECOD	Dictionary-Derived Term	AE.AEDECOD (MedDRA Version 11.1)
AEBODSYS	Body System or Organ Class	AE.AEBODSYS
TRTEMFL	Treatment Emergent Analysis Flag	If ADSL.TRTSDT <= ASTDT<=(ADSL.TRTEDT +14) then TRTEMFL='Y'

General OCCDS Traceability Metadata

- Variable Metadata for ADAE

Name	Variable Label	Variable Metadata
AESTDTC	Start Date/Time of Adverse Event	AE.AESTDTC
AESEV	Severity/Intensity	AE.AESEV
ASEV	Analysis Severity/Intensity	If AE.AESEV='MILD' then ASEV='Mild' Else if AE.AESEV='MODERATE' then ASEV='Moderate' Else if AE.AESEV is equal to 'SEVERE' or Severity/Intensity is missing then ASEV='Severe'
ASTDT	Analysis Start Date	Numeric version by converting AE.AESTDTC from character ISO8601 format to SAS format, applying imputation rules as specified in the SAP or metadata.
ASTDTF	Analysis Start Date Imputation Flag	If start date is completely missing or missing the year then ASTDTF ='Y' Else if start date has month missing then ASTDTF='M' Else if start date has day missing then ASTDTF='D'
TRTSDT	Date of First Exposure to Treatment	ADSL.TRTSDT
TRTEDT	Date of Last Exposure to Treatment	ADSL.TRTEDT
SEX	Sex	ADSL.SEX
RACE	Race	ADSL.RACE

General OCCDS Traceability Output Data

- ADAE Sample Records

Row	STUDYID	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS
1	XYZ	XYZ-001-001	1	HEADACHE	Headache	Nervous system disorders
2	XYZ	XYZ-001-001	2	CHRONIC BACK PAIN	Back pain	Musculoskeletal and connective tissue disorders
3	XYZ	XYZ-001-001	3	NOSE BLEEDING RIGHT NOSTRIL	Epistaxis	Respiratory, thoracic and mediastinal disorders
4	XYZ	XYZ-001-001	4	PROBLEMS OF HYPOTENSION	Hypotension	Vascular disorders
5	XYZ	XYZ-001-001	5	HEADACHE	Headache	Nervous system disorders

Row	AESEV	ASEV	AESTDTC	ASTDT	ASTDTF	TRTEMFL	TRTSDT	TRTEDT	SEX	RACE
1 (cont)	MILD	Mild	2006-01	01JAN2006	D		23JAN2006	15MAY2006	M	ASIAN
2 (cont)	MODERATE	Moderate	2006-01-21	21JAN2006			23JAN2006	15MAY2006	M	ASIAN
3 (cont)	MILD	Mild	2006-01-22	22JAN2006			23JAN2006	15MAY2006	M	ASIAN
4 (cont)	MILD	Mild		23JAN2006	Y	Y	23JAN2006	15MAY2006	M	ASIAN
5 (cont)		Severe	2006-01-24	24JAN2006		Y	23JAN2006	15MAY2006	M	ASIAN

General OCCDS Traceability Summary

- The dataset structure is stated in the dataset metadata
- Dataset structure included
  - data point traceability variables to AE
  - variables from AE needed in analysis or for deriving variables
  - derived variables needed in analysis
  - ADSL covariates needed
- Variable metadata provided the origins of retained variables and derivations of computed variables

Example 2: Traceability When Multiple Input Datasets are Stacked to Create OCCDS

Multiple Input Datasets in OCCDS Analysis Need-1

- The unique –SEQ option from the OCCDS structure allows only one SDTM input for each OCCDS dataset

3.2.2 Identifier Variables

Include the identifier variables from SDTM:

Table 3.2.2.1 OCCDS Identifier Variables

Variable Name	Variable Label	Type	Code List / Controlled Terms	Core	CDISC Notes
STUDYID	Study Identifier	Char		Req	XX.STUDYID
USUBJID	Unique Subject Identifier	Char		Req	XX.USUBJID
SUBJID	Subject Identifier for the Study	Char		Perm	ADSL.SUBJID
SITEID	Study Site Identifier	Char		Perm	ADSL.SITEID
--SEQ	Sequence Number	Num		Req*	XX.--SEQ This would be copied from the SDTM domain XX. This may be missing for derived rows. Required for traceability back to SDTM.

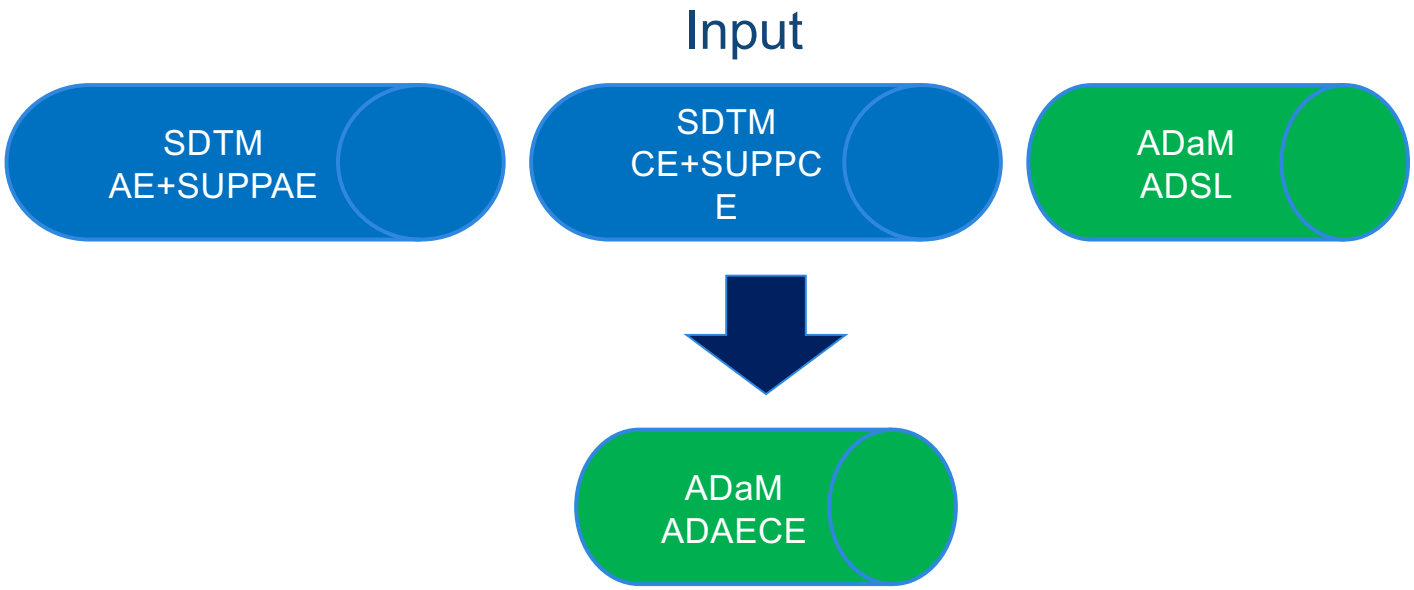
**Note that the only sequence number option shown is --SEQ, because it is unlikely that multiple SDTM domains would be used as input to a single OCCDS dataset.*

- There are instances when it is necessary for multiple SDTM domains to be used as inputs to a single OCCDS dataset

Multiple Input Datasets in OCCDS Analysis Need-2

Summary of All Adverse Events by System Organ Class, MedDRA Preferred Term and Treatment Group Safety Population		
System Organ Class MedDRA Preferred Term	Study Drug + Standard of Care (N=XXX)	Standard of Care (N=XXX)
Subjects with at least one Adverse Event	xx (xx.x)	xx (xx.x)
Blood and lymphatic system disorders	xx (xx.x)	xx (xx.x)
Anaemia deficiencies	xx (xx.x)	xx (xx.x)
Lymphadenopathy	xx (xx.x)	xx (xx.x)

Multiple Input Datasets in OCCDS Data Flow



Multiple Input Datasets in OCCDS Input Data-AE

ROW	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	AESER
1	XYZ	AE	XYZ-001-001	1	FEVER	Pyrexia	General disorders and administration site conditions	N
2	XYZ	AE	XYZ-001-001	2	CHILLS	Chills	General disorders and administration site conditions	N
3	XYZ	AE	XYZ-001-001	3	HEADACHE	Headache	Nervous system disorders	N
4	XYZ	AE	XYZ-001-001	4	LOW NEUTROPHILS	Neutropenias	Blood and lymphatic system disorders	N
5	XYZ	AE	XYZ-001-001	5	DIARRHEA	Diarrhea	Gastrointestinal disorders	N
6	XYZ	AE	XYZ-001-001	6	PNEUMONIA	Pneumonia	Infections and infestations	Y
7	XYZ	AE	XYZ-001-001	7	NAUSEA	Nausea	Gastrointestinal disorders	N

ROW	AESER	AEACN	AEREL	AETOXGR	AESTDTC	AEENDTC
1 (cont)	N	DRUG INTERRUPTED	PROBABLY RELATED	3	2014-02-15T20:15	2014-02-17T05:01
2 (cont)	N	DRUG INTERRUPTED	POSSIBLY RELATED	2	2014-02-15	2014-02-17
3 (cont)	N	DOSE NOT CHANGED	POSSIBLY RELATED	1	2014-02	2014-02-17T20:40
4 (cont)	N	DOSE NOT CHANGED	POSSIBLY RELATED	2	2014-04-14T09:21	2014-06-12T08:30
5 (cont)	N	DOSE NOT CHANGED	PROBABLY RELATED	1	2014-05-15	2014-05-16
6 (cont)	Y	DOSE REDUCED	POSSIBLY RELATED	3	2014-05-13	2014-05-15
7 (cont)	N	DOSE NOT CHANGED	PROBABLY RELATED	1	2014-07-12T14:00	2014-07-13T22:00

Multiple Input Datasets in OCCDS Input Data-CE

ROW	STUDYID	DOMAIN	USUBJID	CESEQ	CETERM	CEDECOD	CECAT
1	XYZ	CE	XYZ-001-001	1	PAIN AT THE INFUSION SITE	Pain	LOCAL INFUSION SITE REACTIONS
2	XYZ	CE	XYZ-001-001	2	REDNESS AT THE INFUSION SITE	Skin erythema	LOCAL INFUSION SITE REACTIONS
3	XYZ	CE	XYZ-001-001	3	SWELLING AT THE INFUSION SITE	Edema peripheral	LOCAL INFUSION SITE REACTIONS
4	XYZ	CE	XYZ-001-001	4	RASH AT THE INFUSION SITE	Rash	LOCAL INFUSION SITE REACTIONS

ROW	CEPRESP	CEOCCUR	CEBODSYS	CEACN	CETOXGR	CESTDTC	CEENDTC
1 (cont)	Y	N	General disorders and administration site conditions				
2 (cont)	Y	Y	Skin and subcutaneous tissue disorders	DOSE NOT CHANGED	2	2014-02-15T10:05	2014-02-15T18:00
3 (cont)	Y	Y	General disorders and administration site conditions	DOSE NOT CHANGED	1	2014-02-15T10:30	2014-02-15T18:35
4 (cont)	Y	N	Skin and subcutaneous tissue disorders				

Multiple Input Datasets in OCCDS ADAECE Dataset Metadata

Dataset	Description	Class	Structure	Keys	Purpose
ADAECE	Adverse/Clinical Events Analysis Dataset	OCCURRENCE DATA STRUCTURE	One record per subject per combined preferred term per start datetime	STUDYID, USUBJID, UDECOD, ASTDTM, ASTDT	Analysis

Multiple Input Datasets in OCCDS ADAECE Data

ROW	STUDYID	USUBJID	SAFFL	TRTA	TRTSDT	TRTSDTM	SRCDOM	SRCSEQ	ACAT1
1	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	CE	2	LOCAL INFUSION SITE REACTIONS
2	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	CE	3	LOCAL INFUSION SITE REACTIONS
3	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	1	ADVERSE EVENTS
4	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	2	ADVERSE EVENTS
5	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	3	ADVERSE EVENTS
6	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	4	ADVERSE EVENTS
7	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	5	ADVERSE EVENTS
8	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	6	ADVERSE EVENTS
9	XYZ	XYZ-001-001	Y	SOC + SD	2014-02-15	2014-02-15T10:05	AE	7	ADVERSE EVENTS

ROW	UTERM	UDECOD	UBODSYS	USTDTC	UENDTC
1 (cont)	REDNESS AT THE INFUSION SITE	Skin erythema	Skin and subcutaneous tissue disorders	2014-02-15T10:05	2014-02-15T18:00
2 (cont)	SWELLING AT THE INFUSION SITE	Edema peripheral	General disorders and administration site conditions	2014-02-15T10:30	2014-02-15T18:35
3 (cont)	FEVER	Pyrexia	General disorders and administration site conditions	2014-02-15T20:15	2014-02-17T05:01
4 (cont)	CHILLS	Chills	General disorders and administration site conditions	2014-02-15	2014-02-17
5 (cont)	HEADACHE	Headache	Nervous system disorders	2014-02	2014-02-17T20:40
6 (cont)	LOW NEUTROPHILS	Neutropenias	Blood and lymphatic system disorders	2014-04-14T09:21	2014-06-12T08:30
7 (cont)	DIARRHEA	Diarrhea	Gastrointestinal disorders	2014-05-15	2014-05-16
8 (cont)	PNEUMONIA	Pneumonia	Infections and infestations	2014-05-13	2014-05-15
9 (cont)	NAUSEA	Nausea	Gastrointestinal disorders	2014-07-12T14:00	2014-07-13T22:00

Row	ASTDT	ASTDTF	ASTDTM	AREL	ARELGR1	ATOXGR	UACN	CEPRESP	TRTEMFL
1 (cont)	2014-02-15		2014-02-15T10:05	Definitely Related	Related	2	DOSE NOT CHANGED	Y	Y
2 (cont)	2014-02-15		2014-02-15T10:30	Definitely Related	Related	1	DOSE NOT CHANGED	Y	Y
3 (cont)	2014-02-15		2014-02-15T20:15	Probably Related	Related	3	DRUG INTERRUPTED		Y
4 (cont)	2014-02-15			Possibly Related	Related	2	DRUG INTERRUPTED		Y
5 (cont)	2014-02-15	D		Possibly Related	Related	1	DOSE NOT CHANGED		Y
6 (cont)	2014-02-14		2014-04-14T09:21	Possibly Related	Related	2	DOSE NOT CHANGED		Y
7 (cont)	2014-05-15			Probably Related	Related	1	DOSE NOT CHANGED		Y
8 (cont)	2014-05-13			Possibly Related	Related	3	DOSE REDUCED		Y
9 (cont)	2014-07-12		2014-07-12T14:00	Probably Related	Related	1	DOSE NOT CHANGED		Y

SRCDOM and SRCSEQ are used to point to the domain and record where the data came from. SRCVAR is not used as in OCCDS, users tend to point back to an entire observation or record

The U* variable allows users to preserve the copy feature yet stack same type data into the same column.
U* can only be used for a direct copy as the 'U' indicate "Unmodified".

Multiple Input Datasets in OCCDS
ADAECE Variable Level Metadata-1

Variable Name	Variable Label	Codelists	Variable Metadata
STUDYID	Study Identifier	XYZ	ADSL.STUDYID
USUBJID	Unique Subject Identifier		ADSL.USUBJID
SAFFL	Safety Population Flag	Y='Yes' 'N'='No'	ADSL.SAFFL
TRTA	Actual Treatment	SOC+SD=Standard of Care + Study Drug SOC=Standard of Care	ADSL.TRT01A
TRTSDT	Date of First Exposure to Treatment	yymmdd10.	ADSL.TRTSDT
TRTSDTM	Datetime of First Exposure to Treatment	datetime20.	ADSL.TRTSDTM
SRCDOM	Source Data		Set to 'AE' if record is from AE dataset. Set to 'CE' if record is from CE dataset.
SRCSEQ	Source Sequence Number		Set to AE.AESEQ if record is from AE dataset. Set to CE.CESEQ if record is from CE dataset.
ACAT1	Analysis Category 1	ADVERSE EVENTS LOCAL INFUSION SITE REACTIONS	If record is from AE then ACAT1='ADVERSE EVENTS' Else ACAT1='LOCAL INFUSION SITE REACTIONS'

Multiple Input Datasets in OCCDS ADAECE Variable Level Metadata-2

Variable Name	Variable Label	Codelists	Variable Metadata
UTERM	Reported Term		AE.AETERM if record is from AE dataset CE.CETERM if record is from CE dataset
UDECOD	Dictionary-Derived Term	MedDRA	AE.AEDECOD if record is from AE dataset CE.CEDECOD if record is from CE dataset
UBODSYS	Body System or Organ Class	MedDRA	AE.AEBODSYS if record is from AE dataset CE.CEBODSYS if record is from CE dataset
USTDTC	Start Date/Time of Event	ISO8601	AE.AESTDTC if record is from AE dataset CE.CEBODSYS if record is from CE dataset
UENDTC	End Date/Time of Event	ISO8601	AE.AEENDTC if record is from AE dataset CE.CEBODSYS if record is from CE dataset
ASTDT	Analysis Start Date	yymmdd10.	<Producer will insert derivation here> For example: Date part of AESTDTC. If full date is present convert to numeric. If Day is missing but year and month correspond with treatment start year and month then set day to the start day of treatment. Otherwise assume the first of the month. If Day and Month are missing but Year corresponds with treatment start year then set month and day to treatment start month and day. Otherwise assume January 1st. If start date is completely missing do not impute.
ASTDTF	Analysis Start Date Imputation Flag	DATEF	If start date has month missing then ASTDTF='M' Else if start date has day missing then ASTDTF='D'
ASTDTM	Analysis Start Date/Time	datetime20.	<Producer will insert derivation here> For example: Convert AESTDTC to a numeric datetime variable
AREL	Analysis Causality		If record is from AE then AREL=AE.AEREL converted to proper case
ARELGR1	Pooled Causality Group 1	Related Not Related	If AREL in('Definitely Related' 'Possibly Related' 'Probably Related') then ARELGR1='Related'. Else if AREL in('Not Related' 'Unlikely Related') then ARELGR1='Not Related'.
UTOXGR	Toxicity Grade	'1'='Grade 1' '2'='Grade 2' '3'='Grade 3' '4'='Grade 4' '5'='Grade 5'	Set to AE.AETOXGR if record is from AE dataset Set to CE.CETOXGR if record is from CE dataset
UACN	Action Taken with Study Treatment		Set to AE.AEACN if record is from AE dataset Set to CE.CEACN if record is from CE dataset
CEPRESP	Clinical Event Pre-Specified	'Y'='Yes'	CE.CEPRESP
TRTEMFL	Treatment Emergent Analysis Flag	'Y'='Yes'	<Producer will insert derivation here> For example: Assume TRTEMFL='Y' unless proven that event is not treatment emergent If both the Start Date/Time of the Adverse Event and Treatment are present and populated and Start Date/Time of Adverse Event is prior to Start Date/Time of Treatment (MISSING<ASTDTM<TRTSDTM) then set TRTEMFL to NULL. If either the Start Date/Time of the Adverse Event or the Start Date/Time of Treatment is missing and both the Start Date of the Adverse Event and Treatment are present and and populated and Start Date of Adverse Event is prior to Start Date of Treatment (MISSING<ASTDT<TRTSDT) then set TRTEMFL to NULL. If Start Date of Adverse Event is missing but End Date/Time or End Date is present and prior to Start Date/Time or Start Date of Treatment then set TRTEMFL to NULL.

Multiple Input Datasets in OCCDS Summary

- When stacking multiple input datasets into one unique OCCDS dataset, it is necessary to utilize the SRCDOM and SRCSEQ variables for traceability
- This example considers input domains of the same general observation class
- The idea can be extended to input datasets of varying observation classes with additional harmonization
- This implementation will be included in OCCDS Version 1.1, which is under review

Example 3: Using an Intermediate Dataset for BDS Traceability

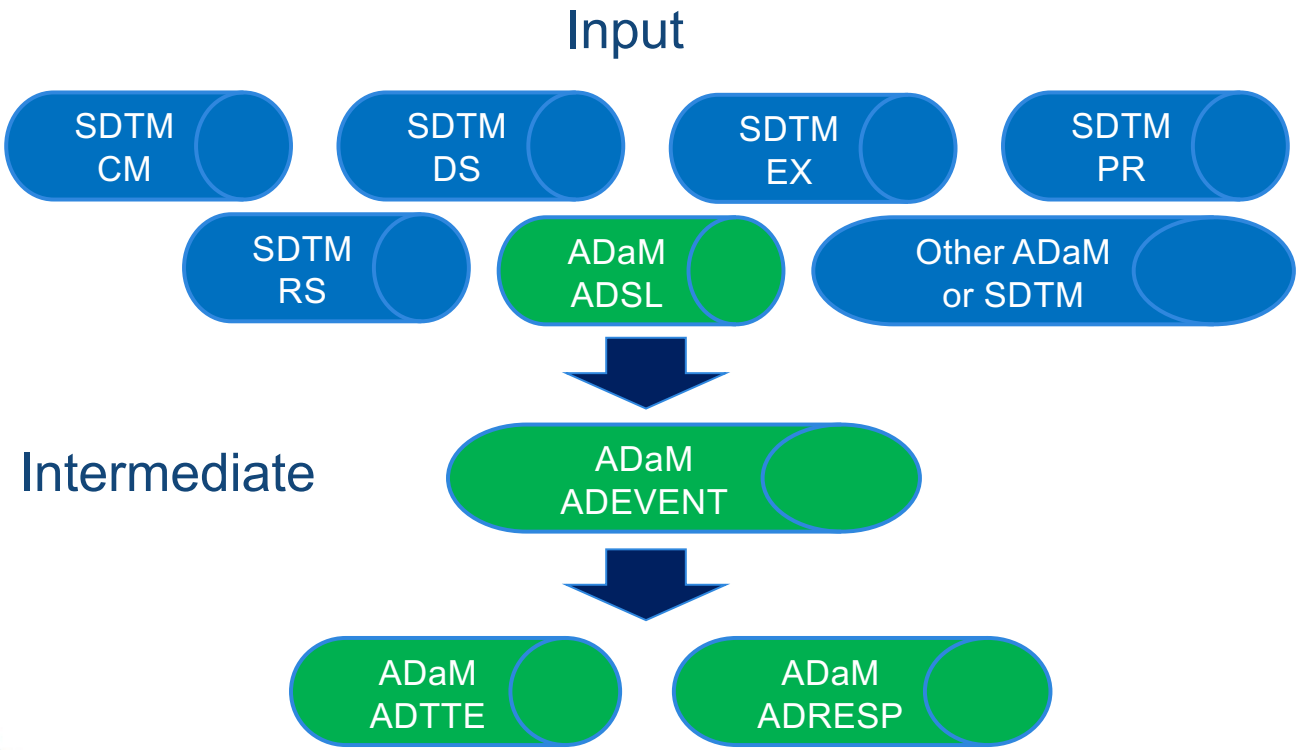
Using an Intermediate Dataset Analysis Need

- Analysis endpoint(s) typically involve complex derivations
- Beneficial to capture all data components used in the derivations in one data set
- Can aid in the review and understanding of the analysis endpoint(s)
- This example expands upon the Breast Cancer Therapeutic Area User Guide (TAUG) section 5.3 example for use of an intermediate event data set to support both time-to-event and best response data sets

Using an Intermediate Dataset Example in the Breast Cancer (TAUG)

- Determine the following time-to-events:
  - Progression Free Survival (PFS)
  - Overall Survival (OS)
  - Duration of Response (DOR)
  - Censor PFS and OS based on censoring rules
- Determine the Best Overall Response (BOR)
- Exclude events that occur after use of a prohibited medication or prohibited procedure

Using an Intermediate Dataset Data Flow



Using an Intermediate Dataset
ADEVENT Intermediate Data

Row	USUBJID	ASEQ	ASTDT	ASTDY	PARAM	PARAMCD	AVALC	SRCDOM	SRCVAR	SRCSEQ
1	ABC-123-001	1	2013-12-29	-4	DISPOSITION	DISPOSIT	RANDOMIZED	DS	DSDECOD	2
2	ABC-123-001	2	2013-12-30	-2	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	RS	RSSTRESC	1
3	ABC-123-001	3	2013-12-31	-1	ASSESSMENT - CENTRAL	ASSESSC	SD	RS	RSSTRESC	2
4	ABC-123-001	4	2014-01-01	1	TREATMENT	TRTM	DRUG A	EX	EXTRT	1
5	ABC-123-001	5	2014-01-21	20	ASSESSMENT - INVESTIGATOR	ASSESSI	SD	RS	RSSTRESC	3
6	ABC-123-001	6	2014-01-22	22	ASSESSMENT - CENTRAL	ASSESSC	SD	RS	RSSTRESC	4
7	ABC-123-001	7	2014-02-13	44	ASSESSMENT - INVESTIGATOR	ASSESSI	PR	RS	RSSTRESC	5
8	ABC-123-001	8	2014-02-14	45	ASSESSMENT - CENTRAL	ASSESSC	PR	RS	RSSTRESC	6
9	ABC-123-001	9	2014-03-06	65	ASSESSMENT - INVESTIGATOR	ASSESSI	PR	RS	RSSTRESC	7
10	ABC-123-001	10	2014-03-07	66	ASSESSMENT - CENTRAL	ASSESSC	PR	RS	RSSTRESC	8
11	ABC-123-001	11	2014-03-28	87	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	RS	RSSTRESC	9
12	ABC-123-001	12	2014-03-29	88	ASSESSMENT - CENTRAL	ASSESSC	PD	RS	RSSTRESC	10
13	ABC-123-001	13	2014-03-30	89	TREATMENT	TRTM	DRUG A	EX	EXTRT	2
14	ABC-123-001	14	2014-03-31	90	EVENT	EVENT	TAMOXIFEN	CM	CMTRT	1
15	ABC-123-002	1	2013-11-10	-3	DISPOSITION	DISPOSIT	RANDOMIZED	DS	DSDECOD	2
16	ABC-123-002	2	2013-11-11	-2	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	RS	RSSTRESC	1
17	ABC-123-002	3	2013-11-12	-1	ASSESSMENT - CENTRAL	ASSESSC	PD	RS	RSSTRESC	2
18	ABC-123-002	4	2013-11-13	1	TREATMENT	TRTM	DRUG B	EX	EXTRT	1
19	ABC-123-002	5	2013-12-01	19	ASSESSMENT - INVESTIGATOR	ASSESSI	SD	RS	RSSTRESC	3
20	ABC-123-002	6	2013-12-02	20	ASSESSMENT - CENTRAL	ASSESSC	SD	RS	RSSTRESC	4
21	ABC-123-002	7	2013-12-14	32	EVENT	EVENT	LUMPECTOMY	PR	PRTRT	1
22	ABC-123-002	8	2013-12-27	45	ASSESSMENT - INVESTIGATOR	ASSESSI	PR	RS	RSSTRESC	5
23	ABC-123-002	9	2013-12-28	46	ASSESSMENT - CENTRAL	ASSESSC	PR	RS	RSSTRESC	6
24	ABC-123-002	10	2013-12-29	47	TREATMENT	TRTM	DRUG B	EX	EXTRT	2

Using an Intermediate Dataset
ADEVENT Variable Level Metadata

ASTDT	Analysis Start Date	Num		The date that the event occurred is the corresponding --DTC variable for each PARAMCD converted to numeric date format. RS.RSDTC when PARAMCD = "ASSESSI" or "ASSESSC" DS.DSSTDTC when PARAMCD = "DISPOSIT" AE.AESTDTC or MH.MHSTDTC or DV.DVSTDTC or CM.CMSTDTC, or PR.PRSTDTC or some other source data for an event which prevents further assessments when PARAMCD = "EVENT".
ASTDY	Analysis Start Relative Day	Num		The number of days from randomization to the date of the reported event. ASTDT – ADSL.RANDDT + 1
PARAM	Parameter	Char	DISPOSITION; ASSESSMENT - INVESTIGATOR; ASSESSMENT - CENTRAL; TREATMENT; EVENT	Set using PARAMCD "DISPOSIT"="DISPOSITION" "ASSESSI"="ASSESSMENT - INVESTIGATOR" "ASSESSC"="ASSESSMENT - CENTRAL" "TRTM"="TREATMENT" "EVENT"="EVENT"
PARAMCD	Parameter Code	Char	DISPOSIT; ASSESSI; ASSESSC; TRTM; EVENT	If RECIST assessment then PARAMCD = "ASSESS" For investigator based tumor response assessments, append "I" to the PARAMCD. For central imaging tumor response assessments, append "C" to the PARAMCD. If disposition event then PARAMCD = "DISPOSIT" If study treatment then PARAMCD = "TRTM" If event that is a protocol violation or prevents further assessments then PARAMCD = "EVENT".
AVALC	Analysis Value (C)	Char		Reported Assessment associated with the ASTDT.
SRCDOM	Source Data	Char		See parameter-level metadata (Table 2.8.3.4)
SRCVAR	Source Variable	Char		See parameter-level metadata (Table 2.8.3.4)
SRCSEQ	Source Sequence Number	Num		The sequence number --SEQ or ASEQ of the row in the dataset identified in the SRCDOM that relates to the analysis value being derived.

Using an Intermediate Dataset
Value Level Metadata

Dataset	Variable Name	Where	Type	Derivation/Comment
ADEVENT	SRCDOM	PARAMCD in ("ASSESSI" "ASSESSC")	Char	Set to "RS"
ADEVENT	SRCDOM	PARAMCD = "DISPOSIT"	Char	Set to "DS"
ADEVENT	SRCDOM	PARAMCD = "TRTM"	Char	Set to "EX"
ADEVENT	SRCDOM	PARAMCD = "EVENT"	Char	Set to "AE", "MH", "DV", "CM" or "PR" based on whether the source of ASTDT is AE.AESTDTC, MH.MHSTDTC, DV.DVSTDTC, CM.CMSTDTC, or PR.PRSTDTC
ADEVENT	SRCVAR	PARAMCD in ("ASSESSI" "ASSESSC")	Char	Set to "AVALC"
ADEVENT	SRCVAR	PARAMCD = "DISPOSIT"	Char	Set to "DSDECOD"
ADEVENT	SRCVAR	PARAMCD = "TRTM"	Char	Set to "EXTRT"
ADEVENT	SRCVAR	PARAMCD = "EVENT"	Char	Set to "AEDECOD", "MHTRT", "DVDECOD", "CMTRT" or "PRTRT" based on whether the source of AVALC is AE.AEDECOD, MH.MHTRT, DV.DVDECOD, CM.CMTRT, or PR.PRTRT

Using an Intermediate Dataset ADTTE and ADRESP Data

ADTTE

Row	USUBJID	PARAM	PARAMCD	AVAL	CNSR	EVNTDESC	SRCSEQ
1	ABC-123-001	Progression-free Survival - Investigator	PFSI	87	0	DOCUMENTED PROGRESSION	11
2	ABC-123-001	Progression-free Survival - Central	PFSC	88	0	DOCUMENTED PROGRESSION	12
3	ABC-123-001	Overall Survival	OS	89	1	CENSORED AT TIME OF LAST ASSESSMENT	13
4	ABC-123-001	Duration of Response	DOR	44	0	DISEASE PROGRESSED	
5	ABC-123-002	Progression-free Survival - Investigator	PFSI	19	1	CENSORED AT TIME OF LAST ASSESSMENT	5
6	ABC-123-002	Progression-free Survival - Central	PFSC	20	1	CENSORED AT TIME OF LAST ASSESSMENT	6
7	ABC-123-002	Overall Survival	OS	1	1	CENSORED AT TIME OF LAST ASSESSMENT	4

ADRESP

Row	STUDYID	USUBJID	PARAM	PARAMCD	AVAL	AVALC	SRCSEQ
1	ABC-123	ABC-123-001	Best Overall Response - Investigator	BORI	2	PR	7
2	ABC-123	ABC-123-001	Best Overall Response - Central	BORC	2	PR	8
3	ABC-123	ABC-123-002	Best Overall Response - Investigator	BORI	3	SD	5
4	ABC-123	ABC-123-002	Best Overall Response - Central	BORC	4	SD	6

Variable Metadata

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Variable Metadata
SRCSEQ	Source Sequence Number	Num		ADEVENT.ASEQ

Using an Intermediate Dataset Summary

- An intermediate data set can
  - Aid in review
  - Help with traceability
  - Ensure the correct record was selected for determining the analysis endpoint(s)
- An intermediate data set is should only contain necessary events needed for the analysis endpoint(s)

Example 4: Traceability When Using a Look-Up Table (LUT)

Traceability When Using a LUT
Analysis Need - 1

- Add variables to data when key variables exist between the data and a LUT
- Classify variables into analysis categories
- This example will show how to identify categories of prohibited medications

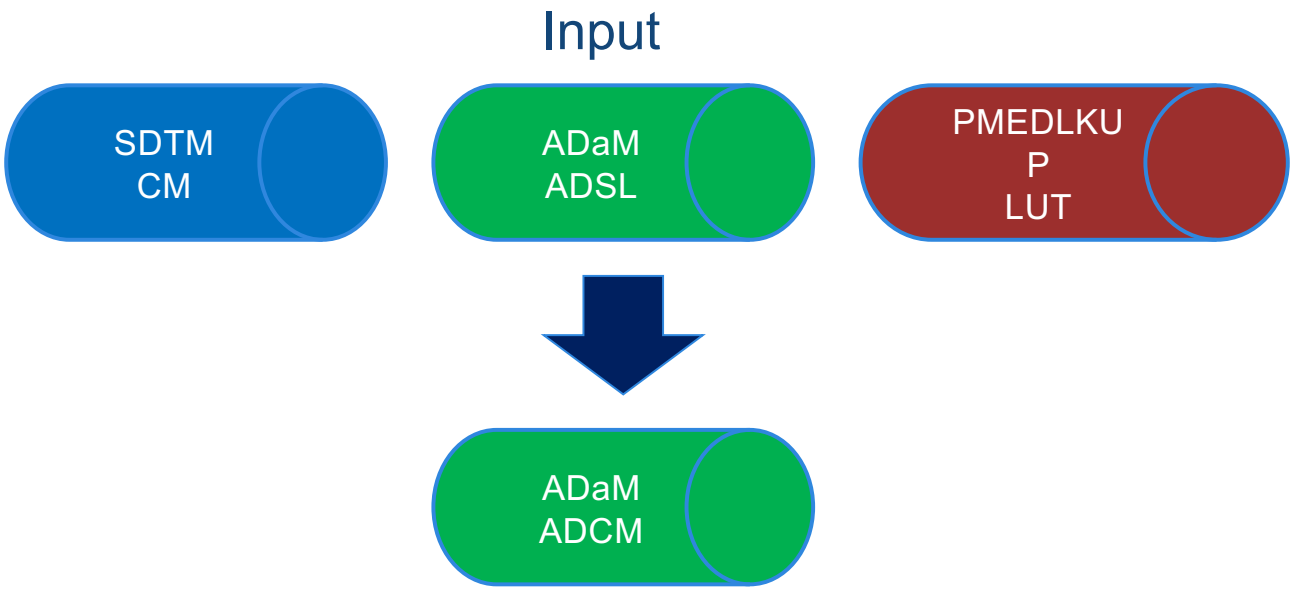
Traceability When Using a LUT Analysis Need - 2

- Sample table shell

Table xx.x
Prohibited Medication During Study
Safety Population

Prohibited Medication Category/ Medication Name	Drug A (N=xx) n (%)	Drug B (N=xx) n (%)	Total (N=xx) n (%)
Category 1	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)
Medication Name 1	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)
Medication Name 2	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)
Category 2	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)
Medication Name 3	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)
Medication Name 4	xx (xxx.x)	xx (xxx.x)	xx (xxx.x)

Traceability When Using a LUT Data Flow



Traceability When Using a LUT
Input Data - CM

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMDECOD	CMCLASCD
1	ABCD	CM	ABCD011001	1	METHYLPREDNISOLONE	H02AB
2	ABCD	CM	ABCD011001	2	ALPHARIX	J07BB
3	ABCD	CM	ABCD011001	3	BECLOMETHASONE	A07EA
4	ABCD	CM	ABCD011002	1	AMOXICILLINE	J01CA
5	ABCD	CM	ABCD011002	2	AZATHIOPRINUM	L04AX
6	ABCD	CM	ABCD021003	1	ADALIMUMAB	L04AB
7	ABCD	CM	ABCD021003	2	PREVENAR	J07AL

Traceability When Using a LUT Look-Up Table - PMEDLKUP

Category	CMCLASCD
Corticosteroid	A01AC
Corticosteroid	D07AC
Corticosteroid	H02AB
Corticosteroid	C05AA
Corticosteroid	A07EA
Corticosteroid	R01AD
Thiopurines	L01BB
Thiopurines	L04AX
Insulins	A10AD

Traceability When Using a LUT
ADCM Variable Level Metadata

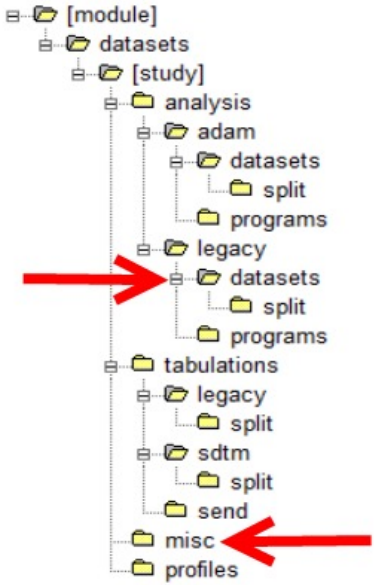
Variable Name	Variable Label	Codelist / Controlled Terms	Variable Metadata
STUDYID	Study Identifier	XYZ	ADSL.STUDYID
DOMAIN	Domain Abbreviation	CM	CM.DOMAIN
USUBJID	Unique Subject Identifier		ADSL.USUBJID
CMSEQ	Sequence Number		CM.CMSEQ
CMDECOD	Standardized Medication Name	Drug Dictionary	CM.CMDECOD
CMCLASCD	Medication Class Code	Drug Dictionary	CM.CMCLASCD
ACAT1	Prohibited Medication Category		Derived: Populate by merging SDTM.CM with the look-up table, PMEDLKUP, by CMCLASCD. Set to the value of CATEGORY from the look-up table if available. Leave as null otherwise.

Traceability When Using a LUT
ADCM Data

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMDECOD	CMCLASCD	ACAT1
1	ABCD	CM	ABCD011001	1	METHYLPREDNISOLONE	H02AB	Corticosteroid
2	ABCD	CM	ABCD011001	2	ALPHARIX	J07BB	
3	ABCD	CM	ABCD011003	3	BECLOMETHASONE	A07EA	
4	ABCD	CM	ABCD011002	1	AMOXICILLINE	J01CA	
5	ABCD	CM	ABCD011002	2	AZATHIOPRINUM	L04AX	Thiopurines
6	ABCD	CM	ABCD021003	1	ADALIMUMAB	L04AB	
7	ABCD	CM	ABCD021003	2	PREVENAR	J07AL	

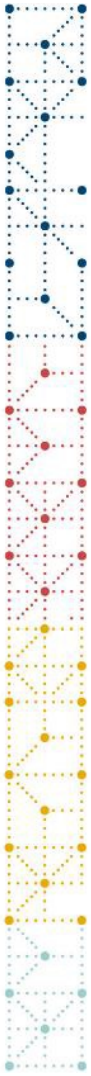
Traceability When Using a LUT Submitting Look-Up Table

- Reference in ADRG
- Provide in MISC folder or LEGACY | DATASETS folder



Traceability When Using a LUT Summary

- Look-Up Tables are a concise mechanism for
  - Applying categories
  - Adding variable values that have a 1 to 1 relationship with collected data
  - Applying consistent PARAM values when PARAMCD values available
- Ensuring documentation in the Analysis Data Reviewer's Guide in the section where it is used helps with traceability



CONCLUSION

Traceability provides transparency to how data is processed into analysis results

Expediates efforts to review, verify and duplicate key endpoints

Comprises of metadata traceability and data-point traceability

Traceability in complex cases may require planning

Good traceability builds trust and confidence in the presented study results

- Special thanks to: Sandra Minjoe, Nancy Brucken, The ADaM Traceability Sub-Team

CONTACT INFORMATION

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

Wayne Zhong Accretion Softworks Wayne321@gmail.com	**Richann Watson** DataRich Consulting richann.watson@datarichconsulting.com
Daphne Ewing CSL Behring Daphne.Ewing@cslbehring.com	**Jasmine Zhang** Boehringer Ingelheim zhangjingwendm@gmail.com

Any brand and product names are trademarks of their respective companies.