



2023

CHINA

INTERCHANGE

BEIJING | 25-26 AUGUST



Insights on Associated Persons Domains

Zhang Xiaofang. Lead Mapping programmer,
Global Data Management and Standards,
MSD R&D (China) Co., Ltd



Meet the Speaker

Zhang Xiaofang

Title: Lead Mapping programmer
Organization: Global Data Management and Standards

- Master degree of Public Health
- Working as mapping programmer for 5 years
- Working in Oncology therapeutic area

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*



Agenda

1. Associated Persons Definition and Application Scenarios
2. Subject Data vs Associated Persons Data
3. Use cases of associated persons dataset



Module 1

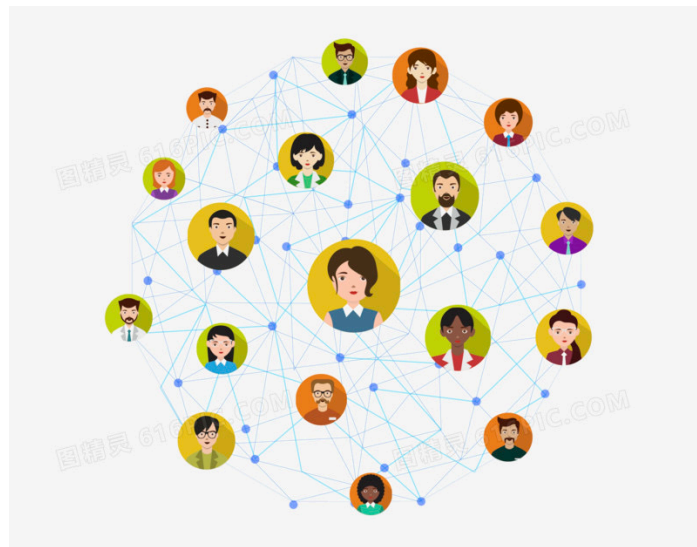
Associated Persons Definition and Application Scenarios

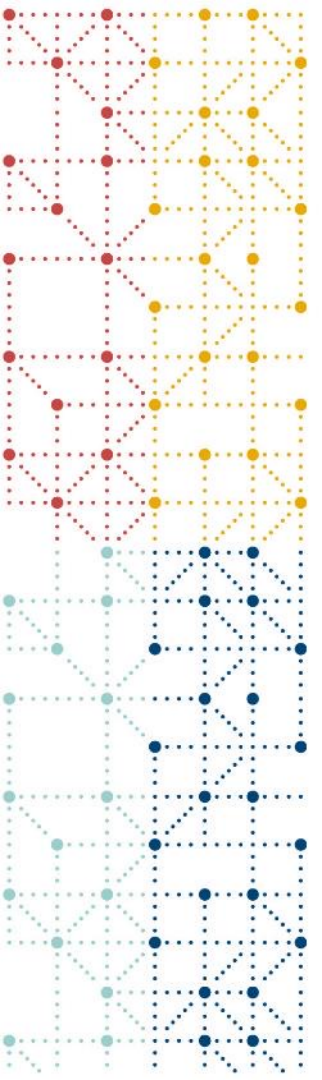
Associated persons

Definition: Associated Persons are not subjects in a clinical study but someone who are associated with the study itself, a particular study subject, or a device used in the study.

Application scenarios:

- Family history
- Donor information
- Caretakers
- Sexual partners
- Environmental exposure
- Diagnostic samples
- Accidental study treatment exposure
- Operator accidents





Module 2

Associated Persons Data VS Subjects Data

Associated person dataset structure

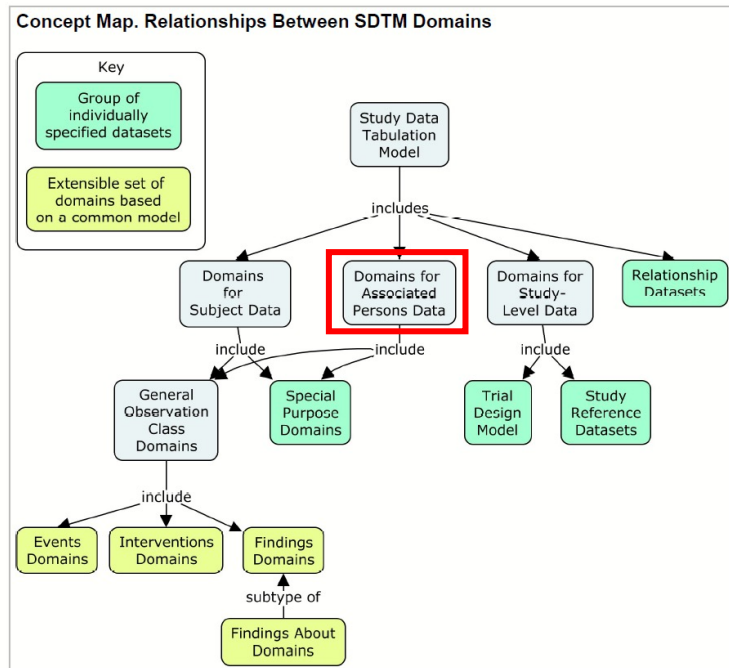
Variable Name	Variable Label	Type	Controlled Terms, Codelist, or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char		Identifier	Four-character abbreviation for the domain, AP--.	Req
APID	Associated Persons Identifier	Char		Identifier <i>Except in APDM where it is the Topic variable</i>	Identifier for a single associated person, a group of associated persons, or a pool of associated persons. If APID identifies a pool, POOLDEF records must exist for each associated person.	Req
--SEQ	Sequence Number	Num		Identifier	Sequence number to ensure uniqueness of records within a domain for associated persons. Unique within APID within the domain.	Req
RSUBJID	Related Subject	Char		Record Qualifier	Identifier for a related study subject or pool of study subjects. The subject(s) may be human or animal. RSUBJID will be populated with the USUBJID of the related subject or the POOLID of the related pool.	Exp
					RSUBJID will be null for data about associated persons who are related to the study but not to any of the study subjects.	
RDEVID	Related Device	Char		Record Qualifier	Identifier for a related device. RDEVID will be populated with the SPDEVID of the related device	Perm
SREL	Subject, Device, or Study Relationship	Char	(RELSUB)	Record Qualifier	If RSUBJID is populated, describes the relationship of the associated person(s) identified in APID to the subject or pool identified in RSUBJID. If RDEVID is populated, describes the relationship of the associated person(s) identified in APID to the device identified in RDEVID. If RSUBJID and RDEVID are null, SREL describes the relationship of the associated person(s) identified in APID to the study identified in STUDYID.	Req

Associated Persons (AP) domain VS Subjects domain

It is necessary to distinguish associated person's data and keep AP data separate from subject data in submission or in data warehouse.

- An AP domain can be created based on traditional SDTM domain.
- Most of the general assumptions about SDTM domains and variables apply to AP data. The structure of AP domains is the same as those for study subject domains.
- AP data is distinguished by the prefix of AP-- in both the domain code and the dataset names.
- Domains SE, SV, DS which describe the progress of a subject through a study are not applicable to Associated persons.
- Supplemental qualifiers for an AP dataset start with a prefix of "SQAP" rather than "SUPP".

SE: Subject Elements SV: Subject Visits DS: Disposition



Associated Persons dataset VS Subjects dataset

- In AP dataset, most of the variable prefixes include only the base domain code
- Associated person datasets include
 - APID (Associated person identifier, rather than USUBJID, a unique subject identifier.)
 - SREL (Describes the relationship of the associated person to the study subject.)
 - RSUBJID (The USUBJID of the study subject with whom they are associated.)
- The following exposure related variables would not generally be used in AP dataset
 RFSTDTTC, RFENDTTC, RFXSTDTTC, RFXENDTTC, RFICDTTC, RFPENDTTC, ARMCD, ARM, ACTARMCD, ACTARM, SITEID

dm.xpt

Row	STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTTC	RFENDTTC	RFXSTDTTC	RFXENDTTC	RFICDTTC	RFPENDTTC	SITEID	INVTNAM	BRTHDTTC	AGE	AGEU	SEX	RACE	ETHNIC	ARMCD	ARM	ACTARMCD	ACTARM	ARMNRS	ACTARMUD	COUNTRY
1	ABC123	DM	ABC12301001	01001	2006-01-12	2006-03-10	2006-01-12	2006-03-10	2006-01-03	2006-04-01	01	JOHNSON, M	1948-12-13	57	YEARS	M	WHITE	HISPANIC OR LATINO	A	Drug A	A	Drug A			USA
2	ABC123	DM	ABC12301002	01002	2006-01-15	2006-02-28	2006-01-15	2006-02-28	2006-01-04	2006-03-26	01	JOHNSON, M	1955-03-22	50	YEARS	M	WHITE	NOT HISPANIC OR LATINO	P	Placebo	P	Placebo			USA

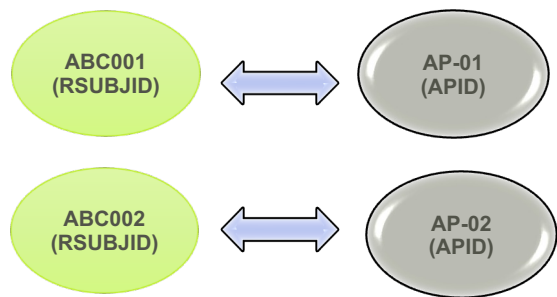
apdm.xpt

Row	STUDYID	DOMAIN	APID	RSUBJID	SREL	INVTNAM	BRTHDTTC	AGE	AGEU	SEX	RACE	ETHNIC	COUNTRY
1	ABC123	APDM	AP-01	ABC12301001	DONOR, ORGAN	JOHNSON, M	1970-07-19	35	YEARS	F	ASIAN	NOT HISPANIC OR LATINO	USA
2	ABC123	APDM	AP-02	ABC12301002	DONOR, ORGAN	JOHNSON, M	1965-09-08	30	YEARS	M	WHITE	HISPANIC OR LATINO	USA

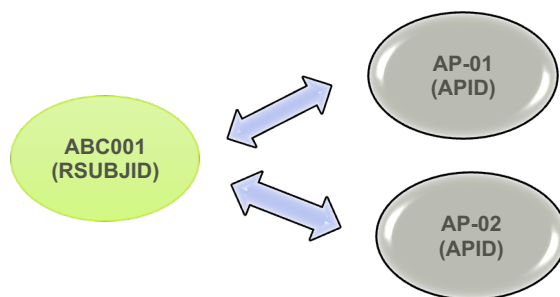
Associated Persons Identifier (APID)

APID --Identifier for a single associated person, a group of associated persons, or a pool of associated person

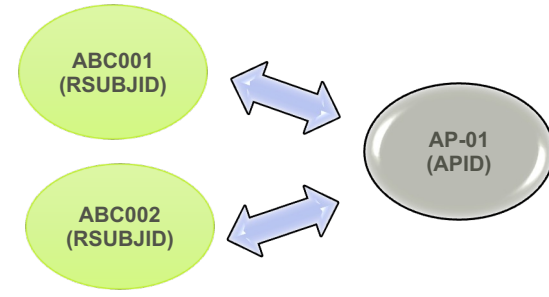
- APID needs to be unique within a study. The same APID should always identify the same associated person.
- When there is no relationship, or the relationship is unknown, distinct APIDs should be used.
- APID is identifier variable except in APDM (It is a Topic variable in APDM).



Subject and AP have 1 to 1 relationship



1 subject has two APs.



1 AP is related to 2 subjects.

RSUBJID and RDEVID

RSUBJID: Populated with the USUBJID of the related subject or the POOLID of the related pool.

- If AP has relationships with multiple subjects, RSUBJID is populated with 'MULTIPLE'.
- If AP is related to study, but not to any study subjects, RSUBJID is populated with NULL.

e.g., A hospital patient, not study subject, is exposed to study treatment due to a clerical error.

apex.xpt

Row	STUDYID	DOMAIN	APID	EXSEQ	RSUBJID	SREL
1	185CUR73	APEX	AP_01	1		ACCIDENTAL ASSOCIATION

Row	EXTRT	EXDOSE	EXDOSU	EXDOSFRM	EXDOSFRQ	EXROUTE	EXSTDTC	EXENDTC
1 (cont.)	CURALL	25	mg	CAPSULE	QD	ORAL	2006-04-01	2006-04-02

RDEVID: Populated with SPDEVID of the related device when devices are study subjects.

e.g., Technician is exposed to radiation when operating study device.

apae.xpt

Row	STUDYID	DOMAIN	APID	AESEQ	RSUBJID	RDEVID	SREL	AETERM	AEDECOD
1	IM_05_2008	APAE	DEV_2011_OP04	1		ABC_123	TECHNICIAN	Operator exposed to radioactive contrast agent	Radiation exposure

Row	AESEV	AESER	AEREL	AEOUT	AESHOSP	AESTDTC	AEENDTC
1 (cont.)	MODERATE	Y	DEFINITELY RELATED	RECOVERING / RESOLVING	Y	2008-04-05	

Subject, Device, or Study Relationship (SREL)

SREL --describes the relationship of the associated person(s) to the subjects, devices or whole study

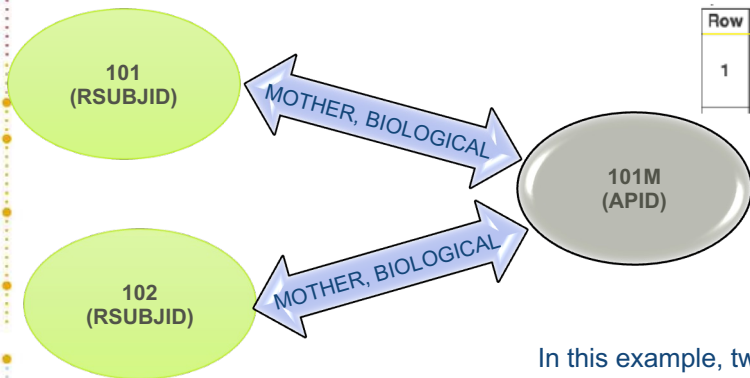
- CDISC codelist Relationship to Subject (RELSUB).
- If an associated person has relationship to multiple subjects or devices and/or multiple relationships to a single subject or device, In those cases, the value 'MULTIPLE' should appear in SREL.
- If value 'MULTIPLE' appear in SREL then APRELSUB dataset is required to save more details.

apce.xpt

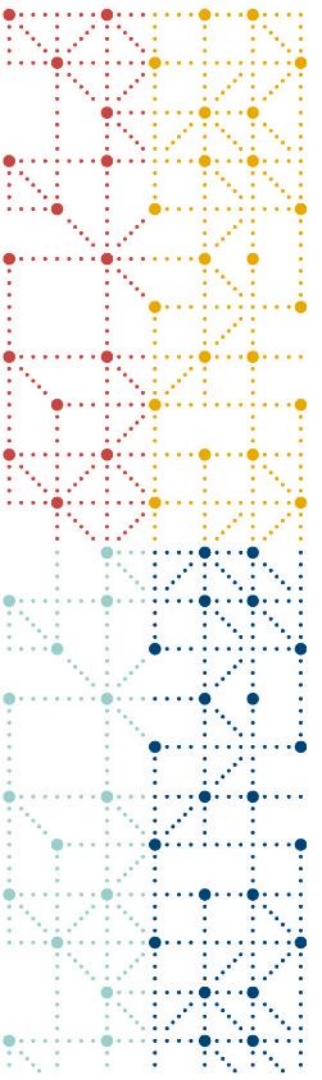
Row	STUDYID	DOMAIN	APID	CESEQ	RSUBJID	SREL	CETERM	CECAT	CESCAT	CESTDTC	CEENDTC
1	XYZ-456	CE	101M	1	MULTIPLE	MULTIPLE	Pregnancy	PREGNANCY-RELATED EVENTS		2017-01-26	2017-10-25

aprebsub.xpt

Row	STUDYID	APID	RSUBJID	SREL
1	XYZ-456	101M	101	MOTHER, BIOLOGICAL
2	XYZ-456	101M	102	MOTHER, BIOLOGICAL



In this example, twins with study identifier 101, 102 are study subjects. Their mother is the associated person, identifier is assigned by appending 'M' to one of the twins' identifier. The mother's pregnancy events are represented in APCE domain.



Module 3

Use cases of associated persons dataset

Use Case#1

A Subject gave birth to twins during study treatment and the sponsor assigned associated person identifiers (APIDs) to infants by appending a letter (“A” and “B”) to the mother’s ID.

Infant Demographics

Identifier:

Date of birth:

Sex: Female Male

Infant Apgar Score

Identifier:

Apgar Score - 1 min:

apdm.xpt

Row	STUDYID	DOMAIN	APID	RSUBJID	SREL	BRTHDTC	SEX
1	ABC-1	APDM	101A	101	SON, BIOLOGICAL	2017-10-25	M
2	ABC-1	APDM	101B	101	DAUGHTER, BIOLOGICAL	2017-10-25	F

aprs.xpt

Row	STUDYID	DOMAIN	APID	RSUBJID	SREL	RSSEQ	RSTEST	RSTESTCD	RSORRES	RSSTRESC	RSSTRESN
1	ABC-1	APRS	101A	101	SON, BIOLOGICAL	1	APGR01-Total Apgar Score	APGR0106	9	9	9
2	ABC-1	APRS	101B	101	DAUGHTER, BIOLOGICAL	2	APGR01-Total Apgar Score	APGR0106	10	10	10

The twins’ data was collected from two different eCRFs. How to ensure APID is consistent between them?

- Directly collect APID from each eCRFs, specify the APID for each infant.
- Collect full APID in one of the eCRFs and indicate associated sequence number in the other and then use the relationship to pull the APID from initial eCRF.
- The same identifier (A, B, et.al.) could be collected from eCRFs to indicate data belongs to which infant, then sponsor use the identifier to derive APID.

USE CASE#2

Family Pompe Disease History is collected about a study subject with the unique identifier of "2011-02-02-031".

Family Pompe Disease History

Have any of the following biological relatives been diagnosed with Pompe Disease?

Mother: Yes No
 Father: Yes No
 Sibling(s): Yes No If yes, how many? _____
 Cousin(s): Yes No If yes, how many? _____

apmh.xpt

Row	STUDYID	DOMAIN	APID	MHSEQ	RSUBJID	SREL	MHTERM	MHPRESP	MHOCCUR
1	2011-02-02	APMH	2011-02-02-N120	1	2011-02-02-031	MOTHER, BIOLOGICAL	POMPE DISEASE	Y	N
2	2011-02-02	APMH	2011-02-02-N121	1	2011-02-02-031	FATHER, BIOLOGICAL	POMPE DISEASE	Y	Y
3	2011-02-02	APMH	2011-02-02-NS122	1	2011-02-02-031	SIBLING, FULL	POMPE DISEASE	Y	N
4	2011-02-02	APMH	2011-02-02-NS123	1	2011-02-02-031	COUSIN, BIOLOGICAL	POMPE DISEASE	Y	Y

sqapmh.xpt

Row	STUDYID	RDOMAIN	APID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
1	2011-02-02	APMH	2011-02-02-NS123	MHSEQ	1	NUMDX	Number Diagnosed	2	CRF	

1. APID can represent an individual or a group of people – even if these people are not considered a “POOL”.
2. SREL value is singular, the number of people in the group is defined in the SQAPMH as it is collected.

USE CASE#3

A family member of a study subject (USUBJID=ABC_123) was exposed to Drug X by accident and rash appeared shortly.

apex.xpt

Row	STUDYID	DOMAIN	APID	EXSEQ	RSUBJID	SREL	EXTRT	EXDOSE	EXDOSTXT	EXDOSU	EXDOSFRM	EXDOSFRQ	EXROUTE	EXSTDTC
1	Y6AOF7	APEX	AP_01	1	ABC_123	FAMILY MEMBER	DRUG X		SPLASH		LOTION	ONCE	TOPICAL	2005-05-05

apae.xpt

Row	STUDYID	DOMAIN	APID	AESEQ	RSUBJID	SREL	AETERM	AEDECOD	AESEV	AESER	AEREL	AEOUT	AESTDTC	AEENDTC
1	Y6AOF7	APAE	AP_01	1	ABC_123	FAMILY MEMBER	Rash	Rash	MODERATE	N	RELATED	RECOVERED / RESOLVED	2005-05-05	2005-05-08

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	APID	RSUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	Y6AOF7	APEX		AP_01	ABC_123	EXTRT	DRUG X		1
2	Y6AOF7	APAE		AP_01	ABC_123	AETERM	Rash		1

1. APID should be same for a particular associated person across different SDTM dataset.
2. RELREC dataset structure for AP data is the same as subject data.

USE CASE#4

A mother had a gemellary pregnancy. One of the fetus was stillborn, the other was born alive.

- If the mother was study subject. Above pregnancy event would be represented in CE dataset.

ce.xpt

Row	STUDYID	DOMAIN	USUBJID	CESEQ	CELNKID	CETERM	CECAT	CESCAT	CESTDTC	CEENDTC
1	ABC-123	CE	101	1	PRG1	Pregnancy	PREGNANCY-RELATED EVENTS		2017-01-10	2017-10-25
2	ABC-123	CE	101	2	PRG1	Stillbirth	PREGNANCY-RELATED EVENTS	PREGNANCY OUTCOME	2017-10-25	
3	ABC-123	CE	101	3	PRG1	Unassisted vaginal delivery	PREGNANCY-RELATED EVENTS		2017-10-25	
4	ABC-123	CE	101	4	PRG1	Live birth	PREGNANCY-RELATED EVENTS	PREGNANCY OUTCOME	2017-10-25	
5	ABC-123	CE	101	5	PRG1	Cesarean delivery	PREGNANCY-RELATED EVENTS		2017-10-25	

- If fetus was the subject and mother was associated person. The event would be saved at APCE dataset.

apce.xpt

Row	STUDYID	DOMAIN	APID	CESEQ	RSUBJID	SREL	CETERM	CECAT	CESCAT	CESTDTC	CEENDTC
1	ABC-12X	APCE	FET-001M	1	FET-001	MOTHER, BIOLOGICAL	Pregnancy	PREGNANCY-RELATED EVENTS		2017-01-10	2017-10-25
2	ABC-12X	APCE	FET-001M	2	FET-001	MOTHER, BIOLOGICAL	Live birth	PREGNANCY-RELATED EVENTS	PREGNANCY OUTCOME	2017-10-25	
3	ABC-12X	APCE	FET-001M	3	FET-001	MOTHER, BIOLOGICAL	Cesarean delivery	PREGNANCY-RELATED EVENTS		2017-10-25	

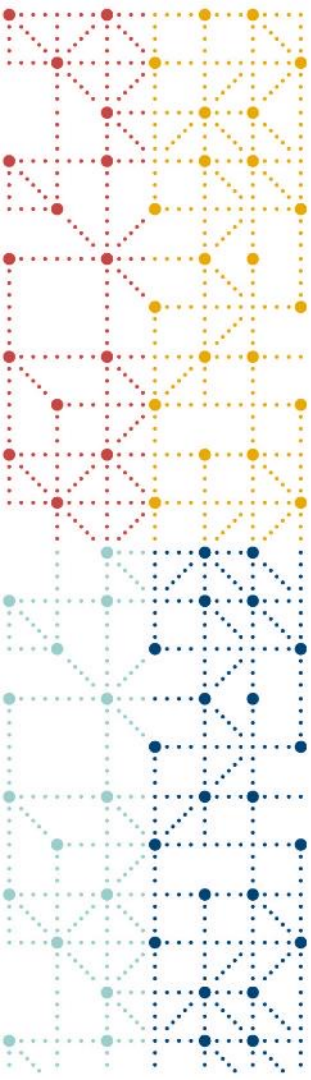
Why the two records in green box above don't present in the APCE dataset?

The two records were about the stillborn infant, who didn't enter the study due to screen failure. Data on screen failures were not submitted, only data related with live-born infant were submitted.



Summary

1. Associated person is someone who associated with the study itself, a particular study subject, or a device used in the study.
2. Most of the SDTM assumptions for subject data are applicable to associated person data.
3. When associated persons' data will be collected from multiple eCRFs, it is critical to ensure APID is same for a particular associated person across different eCRFs.



Thank You!

cdisc