WITH STANDARDS – UNLOCK THE POWER OF DATA

cdisc

2022
US
INTERCHANGE
26-27 OCTOBER | AUSTIN

Sound SDTM, Sound ADaM – Orchestrating SDTM and ADaM Harmonization

Presented by Soumya Rajesh, Senior Standards Engineer, IQVIA and
Nancy Brucken, Senior Standards Engineer, IQVIA

Meet the Speakers

Nancy Brucken

Title: Senior Standards Engineer

Organization: IQVIA

Nancy Brucken has been a veteran CDISC ADaM team member for over 10 years and has contributed to several ADaM publications and sub-teams. She also co-leads the ADQRS sub-team, and is an authorized CDISC instructor, a proud graduate of Marietta College, and a devoted Ohio State Buckeyes fan.

Soumya Rajesh

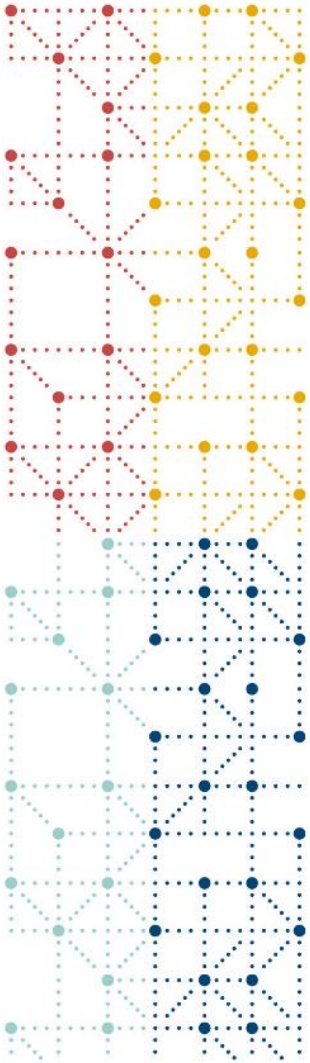
Title: Sr. Standards Engineer

Organization: IQVIA

Soumya Rajesh has 17 years of experience in the areas of SDTM Standards, Programming and Regulatory Operations, in various Therapeutic Areas and study phases. Previous publications cover topics such as Clinical Classifications, Findings About, Disposition, SDTM IG vs. Model, and ISS & ISE Dataset Preparation, at various industry conferences since 2018. Soumya is also the Future Lead for the CDISC SDS LT, Co-Lead of the SDS NSV sub-team, member of CDASH NSV Registry sub-team, PHUSE Working Groups and PharmaSUG Conference Committee.

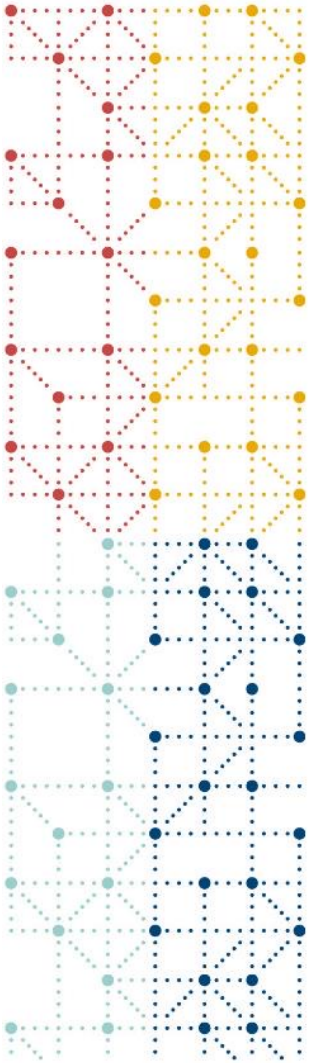
Disclaimer and Disclosures

- *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.*
- *The author(s) have no real or apparent conflicts of interest to report.*



Agenda

1. Introduction
2. Treatment Arms and Elements
3. Reassembling Records After Splitting
4. Disease Response and Clinical Classifications
5. Single or Split FA?
6. --ORRES, --STRESN and –STRESC
7. AE and CM Coding
8. Conclusion



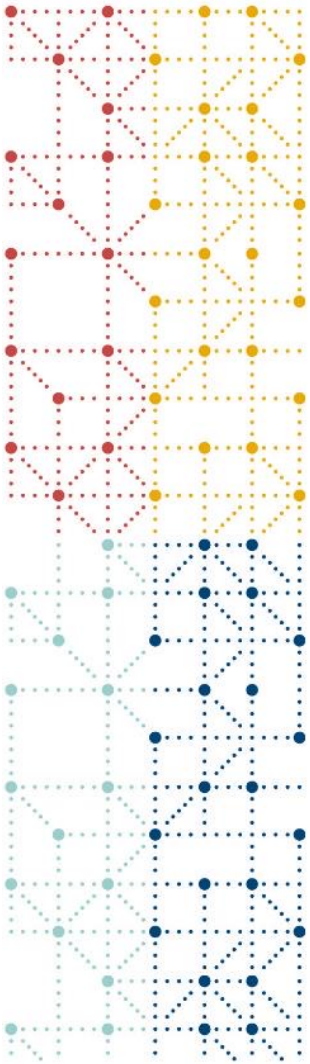
Introduction

Introduction

- Sound SDTM is integral to having sound ADaM
- Data collection forms are designed to be easily completed by sites, investigators and subjects, resulting in higher data quality
- SDTM standards govern the mapping of that data into predictable storage locations, but may result in separation of related pieces of data

How do you balance ease of collection with analysis readiness?

What can you do to make it easier to reassemble the pieces for analysis?



Treatment Arms and Elements

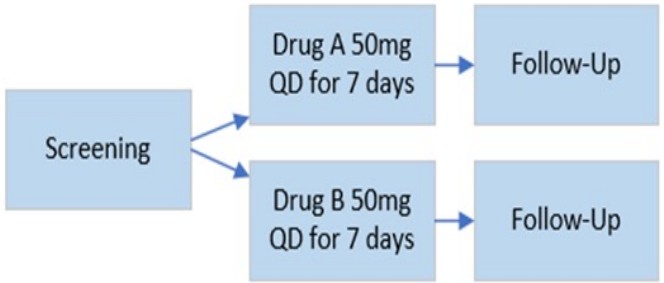
Treatment Arms and Elements contd.

- When creating TA, TE, TV, SE, and SV, focus on the rules and/or assumptions specific to those domains and variables, while balancing the planned trial design, the actual execution of the study, and the study's analyses.
- Study design schema and schedule of assessments can be interpreted in different ways.
- ELEMENT values in TA and TE domains might account for time between doses as rest elements or longer treatment elements with no defined/fixed rest element.
- SDTM should clearly and concisely describe the planned elements while allowing for the actual to be derived and realizing some of this may have reuse in the ADaM ADSL treatment variables.

Treatment Arms and Elements contd.

- For ARMCD, ARM, ETCD, and ELEMENT try to convey some meaning of what is happening.
- ARMCD = ‘1’ conveys no meaning
- ARMCD = ‘A50OL’ identifies the drug, the dose, and that it is open label.
- Especially for pooled analyses:

1. Reuse ARMCD, ARM, ETCD, and ELEMENT values where applicable
2. Distinguish unlike ARM values from each other

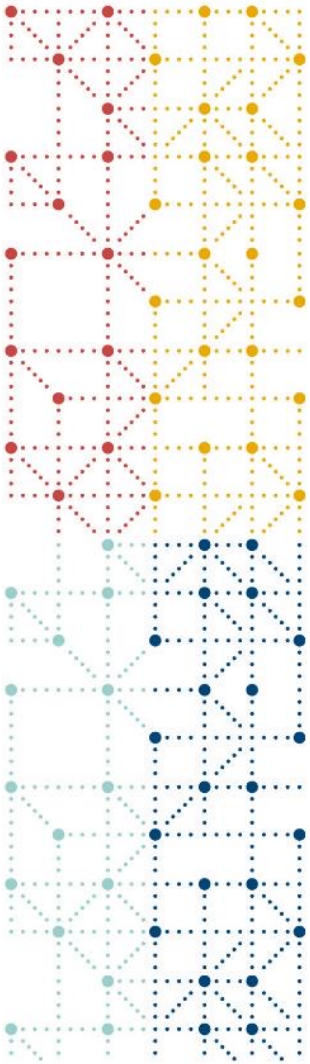


Treatment Arms and Elements contd.

Row	STUDYID	DOMAIN	ARMCD	ARM	TAETORD	ETCD	ELEMENT	TABRANCH	TATRANS	EPOCH
1	ABC-123	TA	A50OL	Drug A 50mg QD P7D Open Label	1	SCRN	Screening	Randomized to Drug A		SCREENING
2	ABC-123	TA	A50OL	Drug A 50mg QD P7D Open Label	2	A50QDOL	Drug A 50mg QD P7D Open Label			OPEN LABEL TREATMENT
3	ABC-123	TA	A50OL	Drug A 50mg QD P7D Open Label	3	FUP	Follow-Up			FOLLOW-UP
4	ABC-123	TA	B50OL	Drug B 50mg QD P7D Open Label	1	SCRN	Screening	Randomized to Drug B		SCREENING
5	ABC-123	TA	B50OL	Drug B 50mg QD P7D Open Label	2	B50QDOL	Drug B 50mg QD P7D Open Label			OPEN LABEL TREATMENT
6	ABC-123	TA	B50OL	Drug B 50mg QD P7D Open Label	3	FUP	Follow-Up			FOLLOW-UP

- Provide clear and accurate descriptions of what is being performed in an element as well as the start rules.
- Elements having no gaps between them in SE also means there should be no gaps between elements in TA and TE.

Row	STUDYID	DOMAIN	ETCD	ELEMENT	TESTRL	TEENRL	TEDUR
1	ABC-123	TE	SCRN	Screening	Informed consent obtained	First dose of study drug	P14D
2	ABC-123	TE	A50QDOL	Drug A 50mg QD P7D Open Label	First dose of drug A	7 days after the start of element or treatment discontinuation	P7D
3	ABC-123	TE	B50QDOL	Drug B 50mg QD P7D Open Label	First dose of drug B	7 days after the start of element or treatment discontinuation	P7D
4	ABC-123	TE	FUP	Follow-Up	End of treatment	14 days after the start of element	P14D



Reassembling Records After Splitting

Reassemble records after splitting

- Related data should be collected and tabulated in a manner that can reflect those relationships (across records or datasets).
- Data that truly should be represented in different domains in SDTM need to be merged back together in ADaM for analysis needs.
- Use dataset or record level relationships and linking variables --LNKID and --LNKGRP (or other variables) where appropriate.
- As an example, best response to a prior anti-cancer medication should not be stored in SUPPCM for convenience's sake. It belongs in the RS domain and a RELREC from the collected prior anti-cancer medications can be established

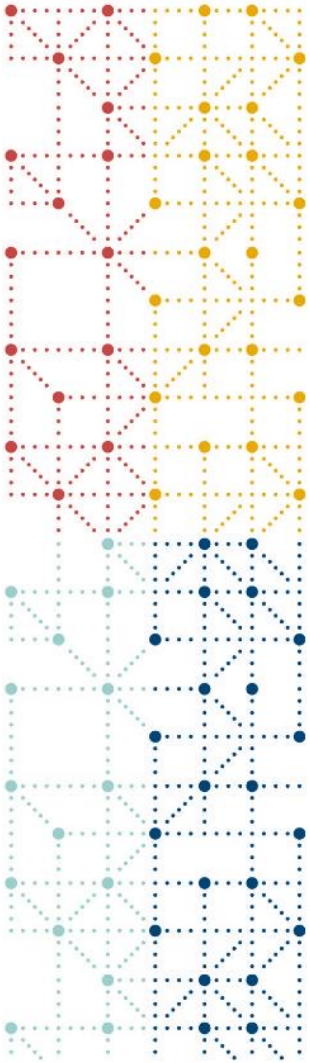
Reassemble records after splitting contd.

- The below example shows how best response to prior anti-cancer medication should mapped in RS and connected to CM with a RELREC by using unique values in CMLNKGRP and RSLNKID that don't conflict with other records in the datasets:

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMGRPID	CMLNKGRP	CMTRT	CMCAT	CMSCAT
1	ABC-123	CM	001-001	1	FOLFOX-001	PRIORAC-FOLFOX-001	FOLFOX	PRIOR ANTI-CANCER	REGIMEN
2	ABC-123	CM	001-001	2	FOLFOX-001	PRIORAC-FOLFOX-001	LEUCOVORIN CALCIUM	PRIOR ANTI-CANCER	CONSTITUENT
3	ABC-123	CM	001-001	3	FOLFOX-001	PRIORAC-FOLFOX-001	FLUOROURACIL	PRIOR ANTI-CANCER	CONSTITUENT
4	ABC-123	CM	001-001	4	FOLFOX-001	PRIORAC-FOLFOX-001	OXALIPLATIN	PRIOR ANTI-CANCER	CONSTITUENT

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSLNKID	RSTESTCD	RSTEST	RSCAT	RSSCAT	RSORRES
1	ABC-123	RS	001-001	1	PRIORAC-FOLFOX-001	BESTRESP	Best Response	RECIST 1.1	PRIOR ANTI-CANCER	PD

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC-123	CM		CMLNKGRP	PRIORAC-FOLFOX-001	MANY	CM-RS-PRIORAC
2	ABC-123	RS		RSLNKID	PRIORAC-FOLFOX-001	ONE	CM-RS-PRIORAC



Disease Response & Clinical Classifications

Disease Response and Clinical Classifications

- The RS domain was initially designed to collect Oncology disease response to treatment. However, it has been expanded to include Clinical Classifications in the SDTMIG v3.3.
- This section shows how to represent historical data, for example - best response in RS.
- With the exception of MH & AE, historical data should be mapped where it belongs, such as CM, LB, VS, EG etc.
- Rather than picking the path of least effort, we need to think about best practices or ideal mapping strategy.
- While the best response to last chemotherapy data is summarized in ADaM, the concern is often how to tie the data from CM or PR to RS.

Disease Response and Clinical Classifications contd.

- RS may include investigator-derived or tool-derived data, but sponsor-derived data should be included in ADaM.
- Oncology response assessments like symptomatic deterioration would often be determined, from TR domain, but could include data from other SDTM domains as well.
- Best response, duration of response, or the progression to prior therapies and follow-up therapies may be represented in the RS domain by linking the record in RS to a group of records or treatment regimens in CM using CMLNKGRP and RSLNKID.

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSLNKID	RSTESTCD	RSTEST	RSCAT	RSSCAT	RSORRES	VISITNUM	RSDTC
1	ABC-123	RS	001-001	1	PRIORAC-FOLFOX-001	BESTRESP	Best Response	RECIST 1.1	PRIOR ANTI-CANCER	PD	1	
2	ABC-123	RS	001-001	1		OVRLRESP	Overall Response	RECIST 1.1	ON-STUDY	CR	12	2010-04-02

Disease Response and Clinical Classifications contd.

- While Best response and Overall response would map under the Disease Response use case, TNM Classification or similar classification / disease staging would map under the Clinical Classification use case of RS.
- Here is an example of how TNM staging would be represented in SDTM based on the initial diagnosis of breast cancer:

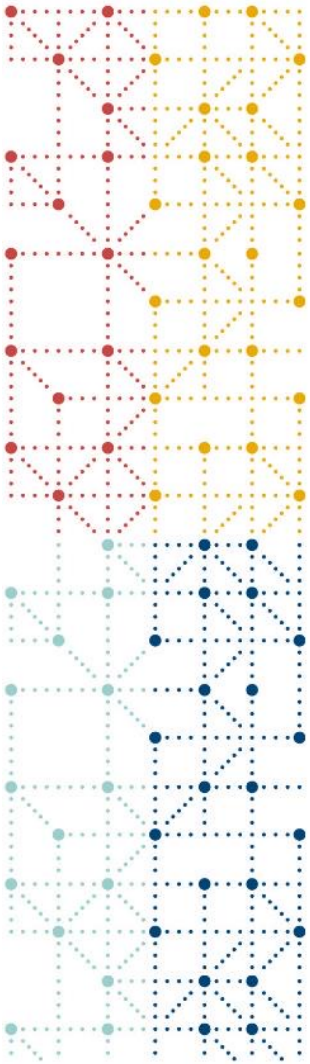
Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHLNKID	MHTERM	MHEVDTYP	MHCAT	MHPRESP	MHOCCUR	MHDTC	MHENRTPT	MHENTPT
1	ABC-123	MH	001-001	1	BR-CA-01	BREAST TUMOR MALIGNANT	INITIAL DIAGNOSIS	PRIOR CANCER HISTORY	Y	Y	2010-10-26	BEFORE	RFSTDTC

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSLNKGRP	RSTESTCD	RSTEST	RSCAT	RSSCAT	RSORRES	VISITNUM	RSDTC
1	ABC-123	RS	001-001	1	BR-CA-01	AJCC101	AJCC1-Primary Tumor (T)	AJCC V7	PRIOR CANCER HISTORY	T1	1	2010-10-26
2	ABC-123	RS	001-001	2	BR-CA-01	AJCC102	AJCC1-Regional Lymph Nodes (N)	AJCC V7	PRIOR CANCER HISTORY	N1	1	2010-10-26
3	ABC-123	RS	001-001	3	BR-CA-01	AJCC103	AJCC1-Distant Metastasis (M)	AJCC V7	PRIOR CANCER HISTORY	M0	1	2010-10-26

Disease Response and Clinical Classifications contd.

- Here is how RELREC would look between MH and RS:

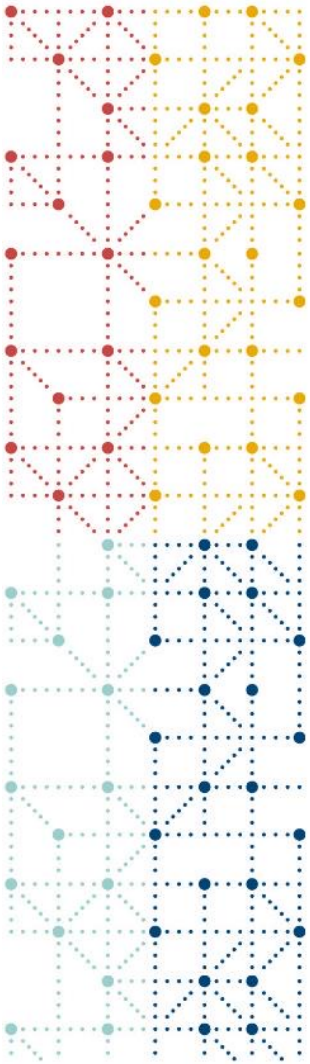
Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC-123	MH		MHLNKID	BR-CA-01	ONE	MH-RS-PRIORAC
2	ABC-123	RS		RSLNKGRP	BR-CA-01	MANY	MH-RS-PRIORAC



SINGLE OR SPLIT FA DATASETS?

Single or split FA datasets?

- Whatever approach you take to represent Findings About data, make sure that it maintains solid traceability.
- Study data pertaining to completely different contexts or parent domains may be better off in split FA datasets.
- If there are complex relationships between FA and multiple parent domains, it may be better to leave the data in a single dataset separated by --CAT, --SCAT and --OBJ, with identifiers like --GRPID / --SPID / --LNKID, to tie the data together by RELREC. Then ADaM can later split this out into multiple analysis datasets as needed.
- Whatever approach you take should consider the analysis or what the ADaM datasets require, and it would all depend on the therapeutic area or the study indication.



--ORRES, --STRESN and --STRESC

--ORRES, --STRESC and --STRESN

- Need a consistent and accurate set of units across subjects and visits for summarizing data and computing change from baseline or other timepoints
- Problems arise most often with lab data, especially when local labs are used
  - Different local labs may report the same test in different units
  - Regulatory agencies may expect certain units for some tests
- If unit conversions are delayed until ADaM, how will regulatory reviewers compare data across subjects?

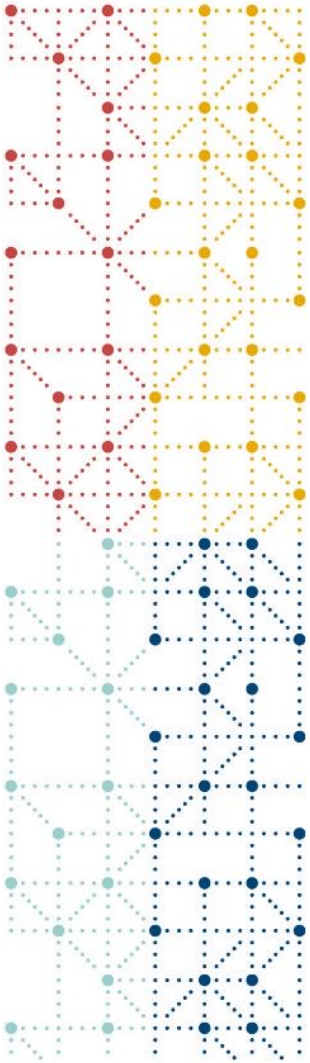
--ORRES, --STRESC and --STRESN

Row	LBTESTCD	LBORRES	LBORRESU	LBSTRESN	LBSTRESU
1	GLUC	5.2	mmol/L	5.2	mmol/L
2	GLUC	80	mg/dL	4.4	mmol/L
3	GLUC	740	mg/L	4.1	mmol/L

- LBORRES holds values in originally collected units
- LBSTRESN holds values converted to a “standard” set of units
- Shows traceability from collected to standardized data and allows reviewers to compare data across subjects
- LBSTRESN may be copied over directly into ADaM AVAL for analysis

QRS/ADQRS Considerations

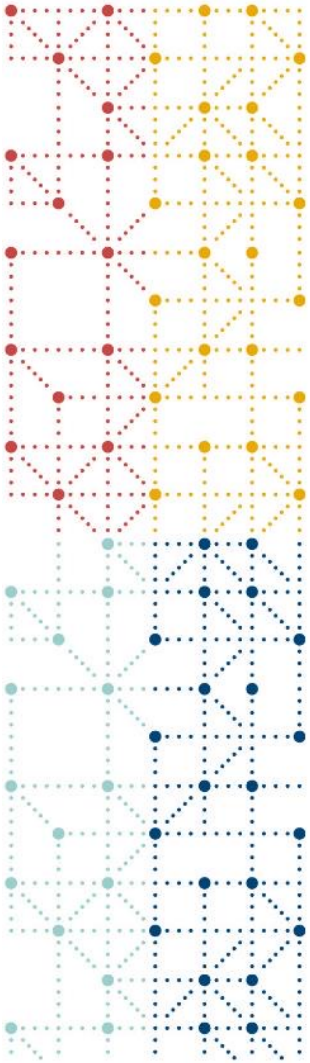
- In SDTM
  - --ORRES contains original text associated with the response
  - --STRESN contains the coded version used for scoring
  - --STRESC contains the character version of --STRESN
- In ADaM
  - AVAL often contains the value from --STRESN
  - AVALC may be populated with the text associated with AVAL if a 1:1 relationship or AVALC is summarized, or --ORRES may be copied over directly for traceability
  - --STRESN may be copied over for traceability if AVAL is different from --STRESN
- If responses need to be reversed or otherwise modified before scoring, need to decide whether to do that in SDTM or ADaM



AE and CM Coding

AE and CM Coding

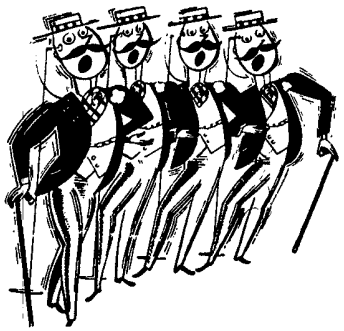
- Variability in how AEs and con meds are reported by sites makes it difficult to summarize collected terms
- MedDRA and WHODrug dictionaries are used to assign collected terms to standard dictionary values
- Coded dictionary terms are copied into ADaM datasets for summarization
- SDTMIG indicates that AEBODSYS and AEDECOD should contain the SOC and preferred term displayed on the AE tables
- CMCLAS and CMDECOD should contain the ATC class and preferred term displayed on the con med tables
- Storing these terms in SDTM provides traceability between SDTM, ADaM and TFLs



Conclusion

Conclusion

- Having sound SDTM is key to having sound ADaM:
  - Use of linking and grouping variables to allow for easy merging of SDTM datasets
  - Creation of RELRECs to explicitly describe relationships between datasets
  - Designing trial design datasets and assigning ARM and ACTARM values with input from biostatisticians and statistical programmers
  - Splitting data in the right manner so its easy to combine later
  - Standardizing measurement units in --STRESN
  - Storing summarized coded variables in SDTM
- Employing these techniques will go a long way towards the creation of harmonized SDTM and ADaM datasets.



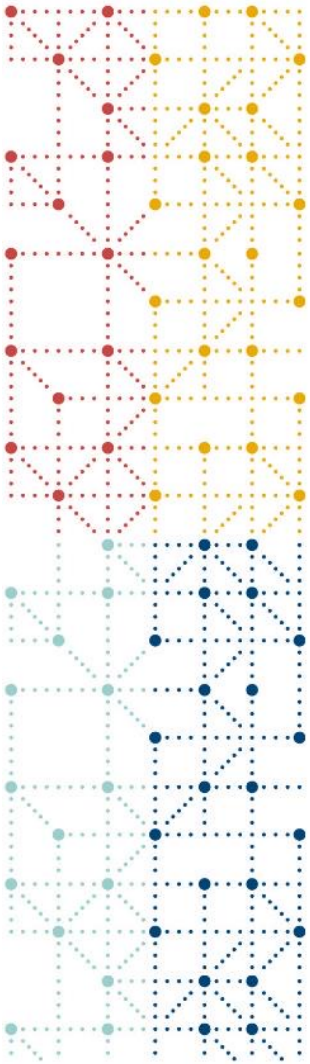
Questions?

Contact Us:

Soumya Rajesh
soumya.rajesh@iqvia.com

Nancy Brucken
nancy.brucken@iqvia.com

David Neubauer
david.neubauer@iqvia.com



Thank You!!