ADaM Pet Peeves: Things Programmers Do That Make Us Crazy

Nancy Brucken, Senior Standards Engineer, IQVIA
Sandra Minjoe, Senior Principal Clinical Data Standards Consultant, ICON PLC
October 13, 2025

Meet the Speakers

Nancy Brucken

Title: Senior Standards Engineer

Organization: IQVIA

Nancy Brucken in a Senior Standards Engineer at IQVIA. She is a CDISC-authorized ADaM instructor, a member of the ADaM 3.0 and Informative Documents sub-teams, and co-leads the ADaM ADQRS sub-team. A graduate of Marietta College, she is a devoted Ohio State fan despite living in that state up north.

Sandra Minjoe

Title: Senior Principal Clinical Data Standards Consultant

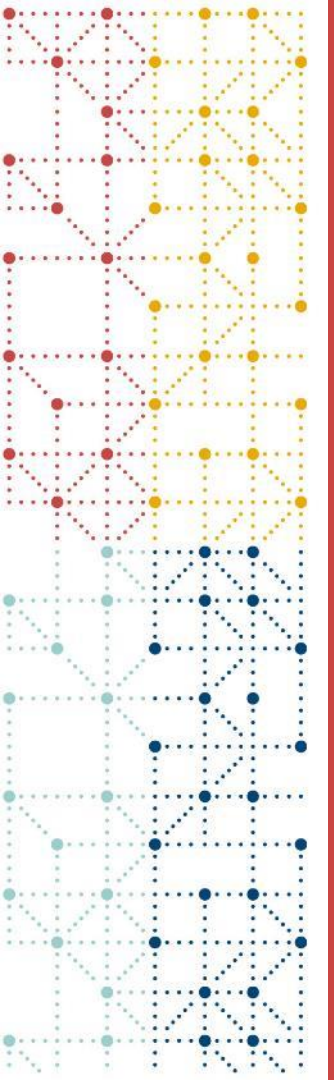
Organization: ICON

Sandra Minjoe is as a Senior Principal Clinical Data Standards Consultant at ICON PLC. She has been part of the CDISC ADaM team since 2001, proposed structures that became ADSL and OCCDS, is a former ADaM Team Lead, is a certified CDISC ADaM trainer, and continues to work on sub-teams.

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

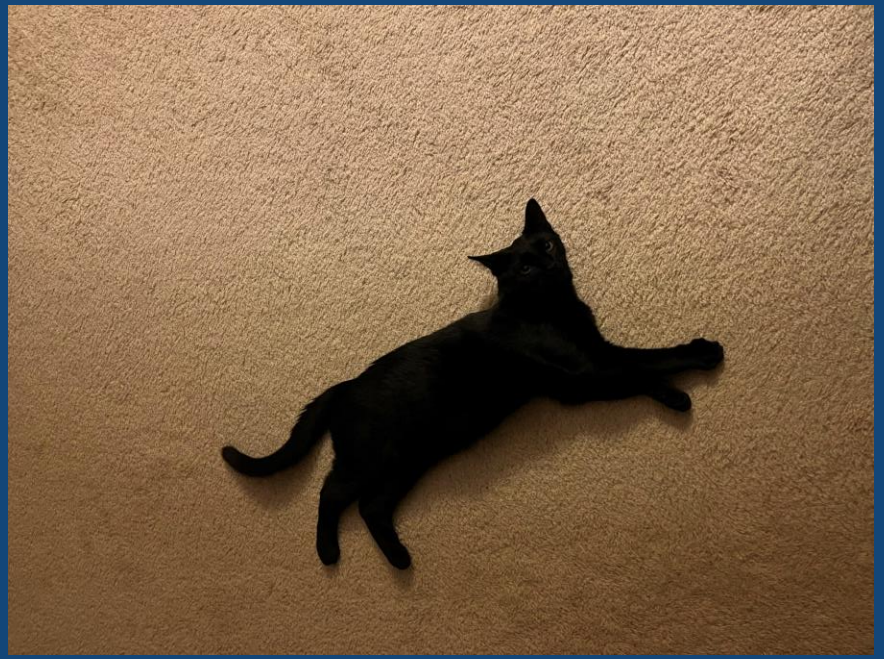
Pet Peeves

1. Intermediate Datasets
2. Use of DTYPE
3. Findings About
4. Parameter Categories and Qualifiers
5. Variable Ordering
6. The Number of ADaM Datasets

For each, we provide a “better” recommended approach

Pet Peeve #1: Intermediate Datasets

(by pet Midnight)



Pet Peeve #1: Intermediate Datasets

- Complex analysis datasets with complicated derivations make it difficult to identify and trace the links back to SDTM
  - The equivalent of not showing your work in a school math problem
- The most complex derivations require the most traceability

Intermediate Datasets Example

- Simple example: Overall Survival (Time to Death)
  - Derived as # days between randomization and death
  - For subjects who have not died, censor at latest of study end date, last assessment date, discontinuation date, …
- More complex: Progression-Free Survival
  - Multiple potential event and censor dates
- Imagine all of these potential dates as separate variables
  - No way to show where each date came from
  - Dates in columns will not be ordered the same across all subjects

Intermediate Datasets Example

- Example Intermediate dataset from Prostate Cancer TAUG

addates.xpt

Row	STUDYID	USUBJID	ASEQ	ADT	ADTDESC	ADTDESCD	ADY
1	ABC-123	ABC-123-001	1	03MAR2014	Date of Randomization	RANDDT	1
2	ABC-123	ABC-123-001	2	15OCT2014	Change in Anti-Cancer Therapy	RXCHGDT	227
3	ABC-123	ABC-123-001	3	15SEP2014	Date of Last Tumor Assessment with No PD	LNOPDDT	197
4	ABC-123	ABC-123-001	4	03DEC2014	Date Last Known Alive	LSTALVDT	276
5	ABC-123	ABC-123-001	5	01NOV2014	Date of Analysis Cut-off	CUTOFFDT	244
6	ABC-123	ABC-123-002	1	16MAY2014	Date of Randomization	RANDDT	1
7	ABC-123	ABC-123-002	2	08JUL2014	Date of Last Tumor Assessment with No PD	LNOPDDT	49
8	ABC-123	ABC-123-002	3	03AUG2014	Date of Tumor Assessment with PD	PDDT	85

One row for each date that could be used as an event or censor

Intermediate Datasets Example

- Example Intermediate dataset from Prostate Cancer TAUG

addates.xpt

Row	STUDYID	USUBJID	ASEQ	ADT	ADTDESC	ADTDESCD	ADY
1	ABC-123	ABC-123-001	1	03MAR2014	Date of Randomization	RANDDT	1
2	ABC-123	ABC-123-001	2	15OCT2014	Change in Anti-Cancer Therapy	RXCHGDT	227
3	ABC-123	ABC-123-001	3	15SEP2014	Date of Last Tumor Assessment with No PD	LNOPDDT	197
4	ABC-123	ABC-123-001	4	03DEC2014	Date Last Known Alive	LSTALVDT	276
5	ABC-123	ABC-123-001	5	01NOV2014	Date of Analysis Cut-off	CUTOFFDT	244
6	ABC-123	ABC-123-002	1	16MAY2014	Date of Randomization	RANDDT	1
7	ABC-123	ABC-123-002	2	08JUL2014	Date of Last Tumor Assessment with No PD	LNOPDDT	49
8	ABC-123	ABC-123-002	3	03AUG2014	Date of Tumor Assessment with PD	PDDT	85

Sortable by date, so a reviewer can easily see the order of events per subject

Intermediate Datasets Example

- Example Intermediate dataset from Prostate Cancer TAUG

addates.xpt

Row	STUDYID	USUBJID	ASEQ	ADT	ADTDESC	ADTDESCD	ADY
1	ABC-123	ABC-123-001	1	03MAR2014	Date of Randomization	RANDDT	1
2	ABC-123	ABC-123-001	2	15OCT2014	Change in Anti-Cancer Therapy	RXCHGDT	227
3	ABC-123	ABC-123-001	3	15SEP2014	Date of Last Tumor Assessment with No PD	LNOPDDT	197
4	ABC-123	ABC-123-001	4	03DEC2014	Date Last Known Alive	LSTALVDT	276
5	ABC-123	ABC-123-001	5	01NOV2014	Date of Analysis Cut-off	CUTOFFDT	244
6	ABC-123	ABC-123-002	1	16MAY2014	Date of Randomization	RANDDT	1
7	ABC-123	ABC-123-002	2	08JUL2014	Date of Last Tumor Assessment with No PD	LNOPDDT	49
8	ABC-123	ABC-123-002	3	03AUG2014	Date of Tumor Assessment with PD	PDDT	85

Not shown: SRCDOM, SRCVAR, and SRCSEQ to point back to input data for each row

Intermediate Datasets Example

- Example Intermediate dataset from Prostate Cancer TAUG

addates.xpt

Row	STUDYID	USUBJID	ASEQ	ADT	ADTDESC	ADTDESCD	ADY
1	ABC-123	ABC-123-001	1	03MAR2014	Date of Randomization	RANDDT	1
2	ABC-123	ABC-123-001	2	15OCT2014	Change in Anti-Cancer Therapy	RXCHGDT	227
3	ABC-123	ABC-123-001	3	15SEP2014	Date of Last Tumor Assessment with No PD	LNOPDDT	197
4	ABC-123	ABC-123-001	4	03DEC2014	Date Last Known Alive	LSTALVDT	276
5	ABC-123	ABC-123-001	5	01NOV2014	D		
6	ABC-123	ABC-123-002	1	16MAY2014			
7	ABC-123	ABC-123-002	2	08JUL2014	Date of Las		
8	ABC-123	ABC-123-002	3	03AUG2014	Date o		

Intermediate dataset ASEQ connects with Time-to-Event dataset SRC* variables

STUDYID	USUBJID	PARAMCD	STARTDT	SRCDOM	SRCSEQ	ADT	AVAL	CNSR
ABC-123	ABC-123-001	PFS	03MAR2014	ADDATES	3	15SEP2014	197	1
ABC-123	ABC-123-002	PFS	16MAY2014	ADDATES	3	03AUG2014	85	0

Recommendation: Intermediate Datasets

- Intermediate datasets can be useful in complex derivations to:
  - Show traceability
  - Enable others to consider alternate approaches
  - Simplify coding
  - Speed up QC

Pet Peeve #2: Use of DTYPE

(by pet Sonoma)

Pet Peeve #2: Use of DTYPE

- The CDISC Notes for variable DTYPE describe 3 common situations where DTYPE is to be populated:
  - A new row is added within a parameter with the analysis value populated based on other rows within the parameter.
  - A new row is added within a parameter with the analysis value populated based on a constant value or data from other subjects.
  - An analysis value (AVAL or AVALC) on an existing record is being replaced with a value based on a pre-specified algorithm.
- The ADaMIG also states
  - "If a parameter is wholly derived, such as a Time-to-Event parameter, then it is a misapplication to populate DTYPE for all records in that parameter because, by definition, all records are derived using the same method."
- CDISC Controlled Terminology exists for DTYPE, and is extensible
  - Note: DTYPE can be up to 200 characters

DTYPE Issue 1: Extending DTYPE when a term already exists

- Some DTYPE Controlled Terminology doesn't sufficiently describe how the row was derived
  - Example: DTYPE = 'AVERAGE' … average of what?
  - But metadata is used to describe exactly what was averaged
- Example: SAP says to add two rows:
  - One that averages AVAL of all post-treatment visits for an Average visit
  - Another that averages AVAL of the last 2 visits for an EndPoint visit
- What should be used for DTYPE?
  - Something like "AVERAGEALL" and "AVERAGE2"? Or "DERIVED"?
  - No! "AVERAGE" should be used for both

DTYPE Issue 1: Extending DTYPE when a term already exists

- How to describe these different averages?
  1. Just use Metadata
  2. Add another variable, in addition to DTYPE

PARAM	VSSEQ	AVISIT	AVISITN	AVAL	DTYPE	DTYPEDTL
Pulse (bpm)	101	Baseline	0	80		
Pulse (bpm)	102	Week 1	1	70		
Pulse (bpm)	103	Week 2	2	60		
Pulse (bpm)		Average	6	70	AVERAGE	All rows averaged
Pulse (bpm)		EndPoint	7	65	AVERAGE	Last 2 rows averaged

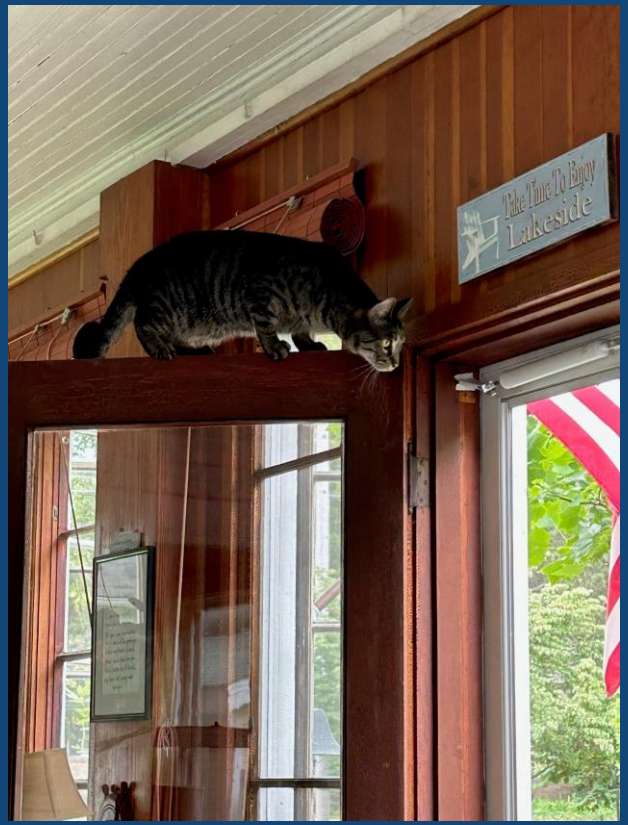
DTYPE Issue 2: Explaining when an entire parameter is derived

- DTYPE is not used to describe how a whole parameter is derived
- Examples of derived parameters:
  - Pediatric study deriving BMI from height and weight
  - Total dose in an exposure dataset

- If you can't use DTYPE to describe detail on how the parameter was derived, what can you do in the data (in addition to metadata)?
  1. PARAMTYP = "DERIVED"
     - PARAMTYP was used in ADaMIG v1.0, deprecated in v1.1
     - Can still be used, and would be a good choice if older studies used it
  2. Add a custom variable such as DPARAMFL ("Derived Parameter Flag")
     - Set to "Y" for all records of a derived parameter

Pet Peeve #3: Findings About

(by pet Kuehner)



Pet Peeve #3: Findings About

- ADaM dataset names must start with letters AD, but can use up to 6 more characters
- Programmers often get in the habit of
  - Naming an ADaM dataset as AD + 2-letter SDTM domain code
    - ADAE, ADEX, ADLB, ADVS, etc.
  - Labeling an ADaM dataset as SDTM label + “Analysis Dataset”

  ... and that works in many circumstances

- But what would you expect to find in a dataset named ADFA with a label “Findings About Analysis Dataset”?

Recommendation: Dataset Name

- Use a name and label that describe what it contains
  - Example: ADMIGRNE with a label of “Migraine Attributes Analysis Dataset”
- Ensure the dataset(s) meet analysis needs
  - FA can contain many different types of data
    - Might need to be split into more than one dataset, each analysis-ready for different sets of tables
    - Some data might be tacked onto other datasets (e.g. AE + FAAE)

Pet Peeve #4: Parameter Categories and Qualifiers

(by pet Thor)

Pet Peeve #5: Parameter Categories and Qualifiers

- The ADaM model and all versions of the ADaMIG state “PARAM must include all descriptive and qualifying information relevant to the analysis purpose of the parameter”
  - Allows analysis and review to make use of a single variable, PARAM (Parameter Name), to describe the result
  - PARAM often includes SDTM TEST content plus qualifiers

PARAM content	SDTM variables needed
Weight (kg)	VSTEST, VSSTRESU
Supine Systolic Blood Pressure (mmHg)	VSPOS, VSTEST, VSSTRESU
Urine Glucose (mg/dL)	LBCAT, LBTEST, LBSTRESU

PARCATy Recommendations

- The same PARAM/PARAMCD value can't belong to more than one PARCATy
  - Example: Urine Glucose and Blood Glucose
    - SDTM uses the same test name with a qualifier
    - ADaM requires unique PARAM/PARAMCD values for each
- Once you've made unique PARAM/PARAMCD values, you can use PARCATy to group them
  - One set of unique parameters for each category (e.g. 'Urine', 'Blood')

PARQUAL History

- Draft ADaMIG v1.2 included variable PARQUAL
  - Not part of final version or in any official ADaM document
  - Included in some TAUGs and an FDA document
- PARQUAL was designed to
  - Hold 1 qualifier (typically something that belongs in the table title)
  - Be used with PARAM
- Only 2 use cases have been defined by the ADaM Team
  - Assessor, such as for Oncology response
  - Drug name/analyte, such as for PK parameters

PARQUAL Recommendations

- OK to use PARQUAL for defined use cases
  - Assessor, such as for Oncology response
  - Drug name/analyte, such as for PK parameters
- Don’t use PARQUAL for anything else
- If using PARQUAL, be prepared to explain conformance issues

- Expect to see PARQUAL from ADaM Team in 2026

Pet Peeve #5: Variable Ordering

(by pet Kuehner)

Pet Peeve #5: Variable Ordering

- ADaM Model:
  - “the ordering of variables follows a logical model (not simply alphabetic)”
  - “It is recommended that the sponsor define a convention for ordering of variables within a dataset and then apply this ordering consistently for all analysis datasets.”
- What order do you use?
  - Common: All ADSL first, then the rest of the variables ordered similar to what is shown the IG
- How will a reviewer easily find what they need?

Recommendation: Variable Ordering

- First: STUDYID, USUBJID
- Second: Key variables used for analysis
  - For BDS, that is probably PARAM/PARAMCD/PARAMN, AVIST/AVISITN, AVAL/AVALC, CHG, etc.
  - For OCCDS, that is probably --TERM/--TRT, dictionary terms
- Third: Traceability variables
- Interleave SDTM counterparts near their ADaM variables
  - VISIT near AVISIT, --STDTC near ASTDT, --STRESN near AVAL, etc.
- Last: end with supportive variables, such as
  - Other dataset-specific variables
  - Other ADSL variables

Pet Peeve #6: Number of ADaM Datasets

(by pet Zoe)

Pet Peeve #6: Number of ADaM Datasets

- Nothing in the ADaM Model or ADaMIG says there must be a 1-1 relationship between SDTM and ADaM datasets
  - OK to combine multiple SDTM datasets into 1 ADaM dataset
    - ADSL contains pieces of multiple SDTM datasets
  - OK to split an SDTM dataset into multiple ADAM datasets for analysis purposes
    - FA split was described in prior slides

Splitting SDTM Data: Examples to Consider

1. QS contains multiple questionnaires
   - ADaM ADQRS subteam recommends splitting by questionnaire

2. LB data needed for both safety and efficacy analyses
   - Safety and efficacy analyses can often include different populations, different visits, and different baseline definitions

3. EG with multiple quantitative and qualitative tests
   - Usually displayed on different tables using different statistical tests

Number of ADaM Datsets: Recommendation

- Consider analysis needs before creating dataset(s)

- Decide what breakdown would be easiest for
  - Deriving analysis values (via variable or value-level metadata)
  - Defining baseline
  - Producing table results
  - Agency review

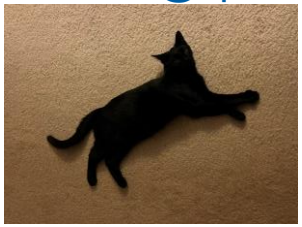
Summary of ADaM Pet Peeves: Things Programmers Do That Make Us Crazy

- Pet Peeves
  1. Intermediate Datasets
  2. Use of DTYPE
  3. Findings About
  4. Parameter Categories and Qualifiers
  5. Variable Ordering
  6. The Number of ADaM Datasets

- For each, we provided a “better” recommended approach

Pet Peeve Questions?

Nancy Brucken
IQVIA
nancy.brucken@iqvia.com



Sandra Minjoe
ICON PLC
Sandra.Minjoe@iconplc.com

Thank You!