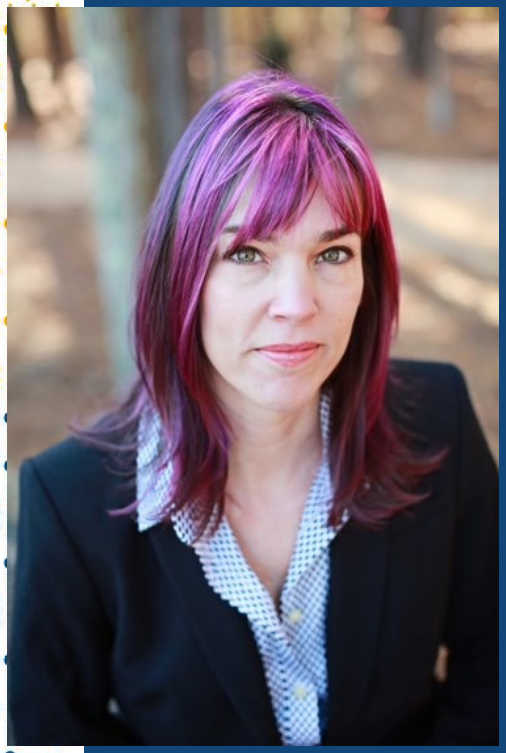
WITH STANDARDS – UNLOCK THE POWER OF DATA

cdisc

2022
US
INTERCHANGE
26-27 OCTOBER | AUSTIN

What is and what should never be: Prevention, Evaluation, and Decisions surrounding source data errors discovered after data lock

Presented by Charity Quick, Director, Statistical Programming and Data Standards, Rho, Inc.



Meet the Speaker

Charity Quick

Title: Director, Statistical Programming and Data Standards

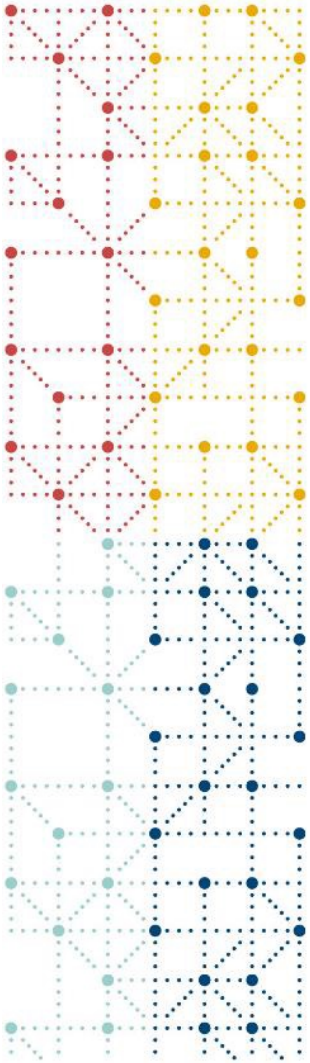
Organization: Rho, Inc.

Charity Quick is the Director of Statistical Programming and Data Standards at Rho – a full service CRO. Her experience involves 15 years of submissions and CDISC conversions for all phases of work as well as managing resourcing, mentoring, and training the Statistical Programming group at Rho. While she is passionate about data standards and loves to engage with others on the topic, she is always eager to talk about music or film instead.

Disclaimer and Disclosures

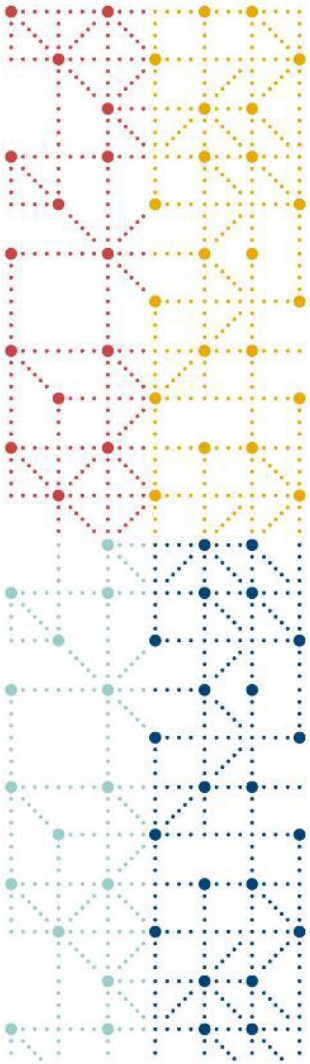
- *The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of CDISC.*

- *The author has no real or apparent conflicts of interest to report.*



Agenda

1. Background
2. Decision Factors – Impact, Perspective, Consultation
3. Documentation
4. Prevention



Background

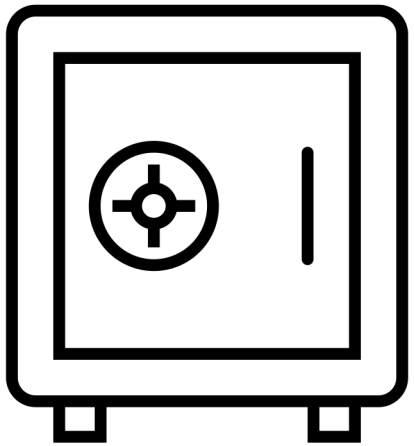
How does this happen and why should we talk about it?

Definitions

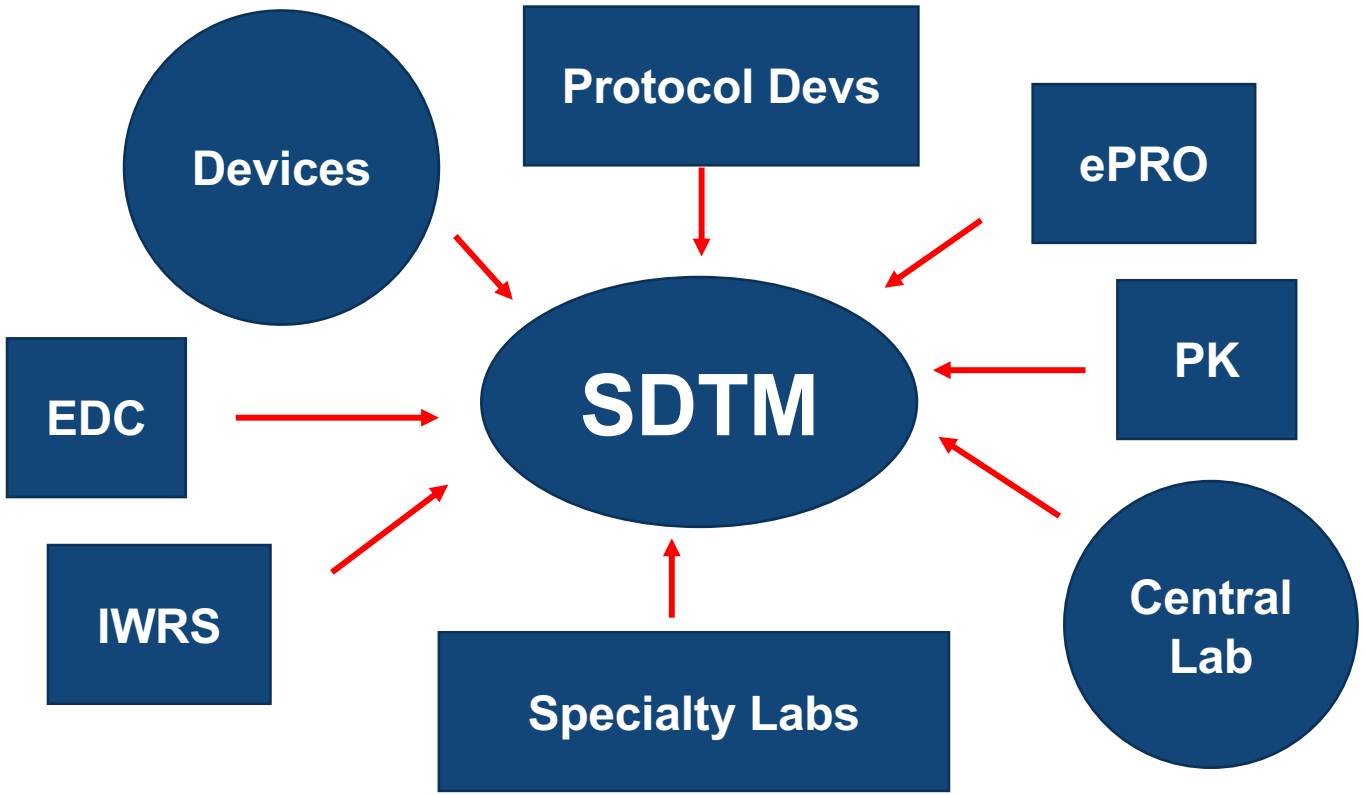
Database Lock – point in time where the clinical trial data is complete and accurate, and no further changes will occur. This includes all trial data as collected.

Source/Raw Data Issue – collected subject data point that is incorrect, inconsistent, missing, or otherwise inaccurate.

Hard Code – correction to inaccurate data by embedding the updated data point directly into the output or otherwise altering source data programmatically.



Data from Everywhere! – increasingly complex



Background

Increasingly High Probability of Data Issues at lock

- More sources, more Vendors, more Handling
- Sometimes inflexible data from vendors
- Pressure to meet database lock timing goal

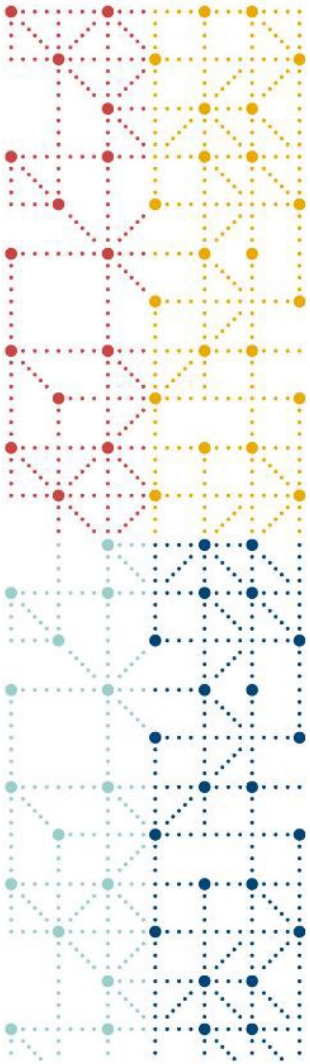
High Quality and Conformance Expectations for study data

- Conformance report and explanations sent to review agencies with CDISC submissions

Background

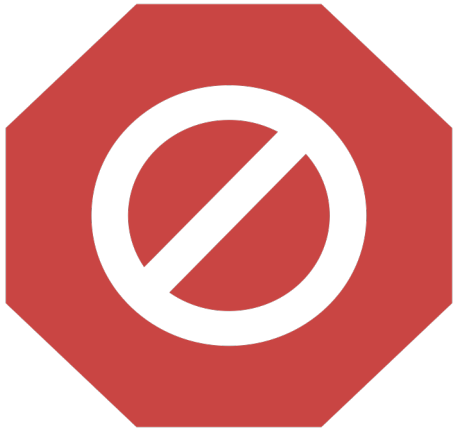
Abundance of Guidance to use

- Internal SOPs and Guides
- Implementation Guides for CDISC deliverables
- Industry Best Practices



Decision Factors

Assess Impact, Maintain Perspective, Consult Experts



Data Issue? Or Mapping Issue?

Assess Study and Data Impact Factors

Study Impact Factors

- What phase is the study?
- Is the study pre-CDISC standards? A legacy conversion?
- Will this study be potentially included in an ISS/ISE?
- Did the study meet endpoint criteria or fail?

Problem Data Impact Factors

- Do the problem data relate to or impact endpoints, SAEs, or other highly important portions of analysis or conformance?

Assess Study and Data Impact Factors - Example

SD0009	No qualifiers set to 'Y', when AE is Serious

Is this an SAE record?

Is this an AE of Special Interest?

Is the program moving towards an ISS?

Maintain Perspective by using Guidance

The SDTM model is a structured formatting of raw data and should provide an integrated data summary of what happened with the subjects on study.

Know the published automatic rejection criteria for the agency you would be using for your study submission

Consult the agency conformance guides and published validation rules

Leverage the issue IMPACT guidance of conformance tools

Think like a reviewer – or try to

Maintain Perspective - Example

SD0002	NULL value in AEDECOD variable marked as Required	Reject
SD0080	AE start date is after the latest Disposition date	High
SD1201	Duplicate records in DV domain	Medium
SD1124	Missing value for PEREASND, when PESTAT is 'NOT DONE'	Low

Maintain Perspective - Example

Sponsors should not add leading or trailing spaces to the USUBJID variable in any dataset. For example, applications have been previously submitted in which the USUBJID variable for each individual subject appeared to be the same across datasets; however, in certain datasets, the actual entry had leading zeros added, or zeros added elsewhere in the entry. This does not allow for machine-readable matching of individual subject data across all datasets. Improper implementation of the USUBJID variable is a common error with applications and often requires sponsors to re-submit their data

From the Study Data Technical Conformance Guide
March 2022
Section 4.1.1.2

Leverage Experience and Consult Others

Seek guidance Internally

- Experiences of other Programmers, Biostatisticians, and SMEs
- Medical, Clinical, Project Management, and Regulatory advice and insight

Consult the Sponsor (if you are CRO/Vendor)

If appropriate, consider emailing the agency

- edata@fda.hhs.gov

Common Decision Outcomes

Unlock EDC or other source data to make corrections

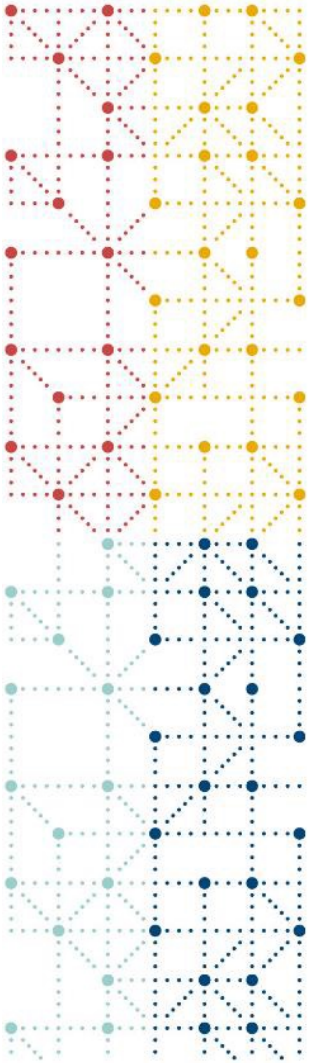
- Source will be consistent with CDISC and analysis

Hard Code Correction in SDTM

- Last resort
- Often used when the incorrect data is in a source that can't be updated or unlocking the EDC is less desirable
- Adequate documentation is necessary

Leave the data point

- Adequate documentation is necessary



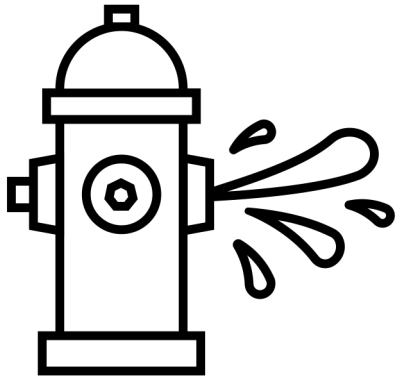
Documentation

An Opportunity for Transparency!

Example Issue for Reference – Decision to hard code

Dosing Diary ePro has incorrect date for **first** dose for 2 subjects.

ePRO vendor programming caused "Evening" doses after midnight to be backdated by one day. 2 subjects showed a first dose PRIOR to randomization in the ePRO data. Vendor could not update.



define.xml

define.xml documentation appropriate for decisions that include hard codes or leaving incorrect data points to pass into CDISC data

- Utilize the Methods section
- Utilize the Comments section
- Consider a link from the csdrg/adrg to the matching define

define.xml – Hard Code Example

erms at	Origin / Source / Method / Comment
	Protocol Set to “SAMPLE01”
	Assigned
	Derived Set to concatenation of STUDYID and SUBJID separated by a "-".
	Derived First non-missing dose date and time as recorded in subject diary. Programmatic correction to device data for 2 subjects Study Data Reviewer’s Guide [12]

Reviewer's Guides

csdg/adrg documentation appropriate for decisions that include hard codes or leaving incorrect data points to pass into CDISC data

- Add information to the impacted dataset section
- Consider adding an appendix that contains a signed NTF (note to file) with information about incorrect data or data that has been hard coded
- Add detail to conformance explanations
- Start collecting notes for your reviewer's guides early while on study
- Utilize an internal review by a team member that has less knowledge of the study data so they can point out places where additional content would be helpful

Reviewer's Guides

Hard Code Example

3.3.10 EX – Exposure

All subject dosing data was collected using an eDiary device. Per protocol, dosing was to occur each morning and evening with the exception of the first dose date (evening only) and the last dose date (morning only). Two study subjects ("XXX-01-005" and "XXX-02-012") had the incorrect dose date logged in their device due to device programming rules that caused their dosing dates to be adjusted incorrectly by one day. Since the incorrect doses were first doses used to determine overall treatment duration, treatment emergent adverse events, etc, the study team decided to correct the data in the SDTM. A programmatic hard code was included in the SDTM EX and DM domains to correct the dosing date error for these 2 subjects. See Appendix II.

Leave Incorrect Data Example

3.4.1 AE - Adverse Events

Adverse Events were coded with MedDRA version 24.0. Subject 100-180 had no end date (AEENDTC) recorded for their "Nausea" even though the outcome (AEOUT) indicated that the event had resolved. See Conformance Details for more information.

Reviewer's Guides – Hard Code Example

Reviewer's Guides – Hard Code Example

Appendix II: Exposure Dates Prior to Randomization/Study Drug Dispensation NTF

To: Sample01 Study Team
CC: Person, Project Manager, VendorX
From: Person, Data Manager, VendorX
Date: 15 Dec 2021
Re: **SAMPLE01 First Dose Prior to Drug Dispensed DIARYpro Data**

This Note to File for SAMPLE01, specifically the data file transferred on 17 NOV 2021, is to explain the occurrences of subject's indicating first dose prior to drug dispense. Two subjects were identified to enter in error times of study medication intake that occurred just past midnight that due to VendorX's remapping correction code, moved the following subject records to a date prior to First Dose:

XXX-01-05: evening dose entered as 12:06am on 04/11/2021. Study drug was dispensed on 04/10/2021. Since the subject labelled this an evening dose and entered as just past midnight on the morning report on the 11th then the device remapping moved it back a day to the 10th from the 11th.

Display footnotes – Leave Incorrect Data Example

Consider adding information about problematic, errant, or otherwise noteworthy data points to TLF footnotes where applicable

11035/ Pbo/Active	61/ F/ B	N/ Y	2018-06-19	DB	2018-07-17	2018-10-09/ 85		Yes		
				DC	2018-10-10/ 86	2018-10-22/ 98		No/		Yes
				OL	2018-10-23/ 99	2019-10-22/ 463/ 365	2019-11-19/ 491/ 393	Yes		Yes

Data Cut: Final Locked Data 2021-04-30
Note: Subject 11035 discontinued during the Discontinuation Period, but has a missing Reason for Discontinuation. This subject was not included in the Reason for Discontinuation summary of the Treatment Discontinuation Period.

Comments in the data creation program

Program documentation appropriate for decisions that include hard codes

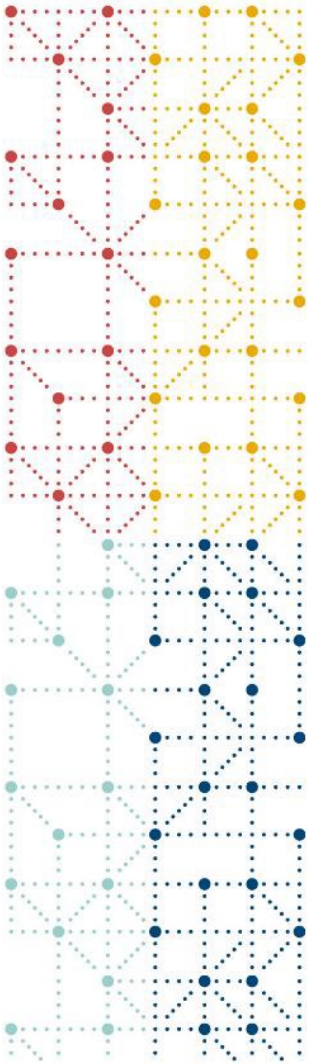
- The code used to create your data is the ultimate documentation
- Comment non-standard items in your code
- Comment and date extra information about hard codes
- Comments have a practical use as well. If that program is copied for use in another study, it's likely that the specialized code should probably be removed so having it commented makes it easier to spot and extract.

Comments in the data creation program – Hard Code Example

```
   PROGRAM HISTORY:

   DATE         PROGRAMMER         DESCRIPTION
   ---------    ---------------    ------------------------------------------------------
   2021-04-12   Charity Quick      create
   2021-09-14   Charity Quick      Update EXDOSU and EXROUTE for CDISC CT
   2021-12-18   Charity Quick      POST-LOCK HARD CODE TO CORRECT DOSE DATE-2 SUBJECTS
*-------------------------------------------------------------------------------------*/

         ************************************************************;
         * HARD CODE 2021-12-18  - SEE EMAIL FROM VENDOR and SPONSOR ;
         * DATED 2021-12-10 and refer to csdrg for additional details;
         ************************************************************;

            if usubijd = "XXX-01-005" and EXSTDTC = "2021-04-09" then
               EXSTDTC = "2021-04-10";
            if usubijd = "XXX-02-012" and EXSTDTC = "2021-10-15" then
               EXSTDTC = "2021-10-16";
```

eTMF documentation

Consider adding documentation to the eTMF if applicable based on the issue, decision, and documentation. Consult your project manager, project lead, and eTMF specialist to determine if eTMF documentation is appropriate.



Prevention

Tips for avoiding post lock data issues

Proactive Issue handling during raw data use

When using data to create SDTM and ADaM run conformance early and often

- Many post-lock issues could have been caught by an early conformance run if the issue was in the raw data prior to lock
- Consider directly involving Data Management in CDISC conformance checks

Document data issues that are seen on study

- Use the Conformance tool interface
- Use a separate data tracker

Invest in Process Improvement

When post lock data issues occur, consider any processes that could be improved to avoid further issues internally and externally

- For example – Protocol Deviations are a common source of post lock data issues
- Suggest a process guide and/or re-training if process gaps are uncovered when investigating a post lock data issue origin

Data Stewardship

Who is responsible for data in your organization?

- Example – if a local lab unit is reported incorrectly who is responsible for identifying the issue before database lock?
- Who reviews Data Transfer Agreements (DTAs) from vendors?

Did the post-lock data issue identify gaps in data stewardship?

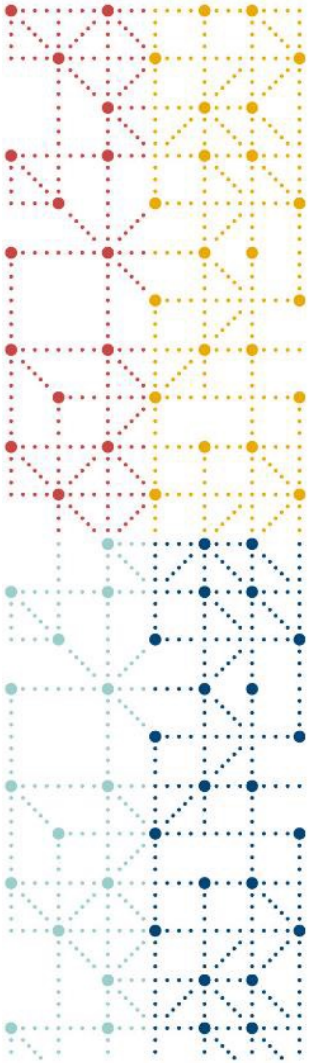
Are efforts being duplicated that could be applied elsewhere?

Encourage Cross-Functional Data Review

Cross Functional CRF review prior to CRF finalization

Suggest Strategically Timed Cross Functional Data Review meetings

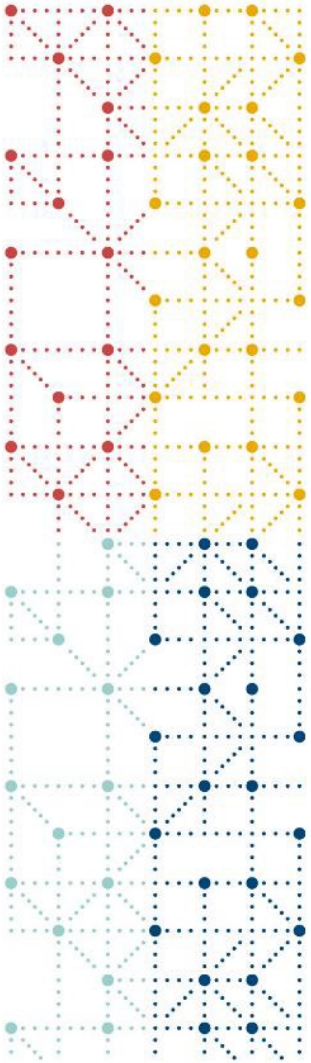
- If possible and depending on enrollment projections, first one after 50% of subjects complete
- Second ~6 weeks prior to lock if timeline permits
- Focus in and across data streams rather than table cosmetics



Summary

1. Incorrect Post-Lock data is likely to happen
2. Assess Impact
3. Maintain perspective by using Guidance and Best Practices
4. Consult other Experts
5. Determine the best decision
6. Document Document Document
7. Prevention and Lessons Learned

8. Questions?



Thank You!

charity_quick@rhoworld.com

cdisc