

Data Transformation: Achieving Compliance with Irregular Data

Eduardo Elizalde, Project Data Manager

Reyes.elizaldeduarte@syneoshealth.com

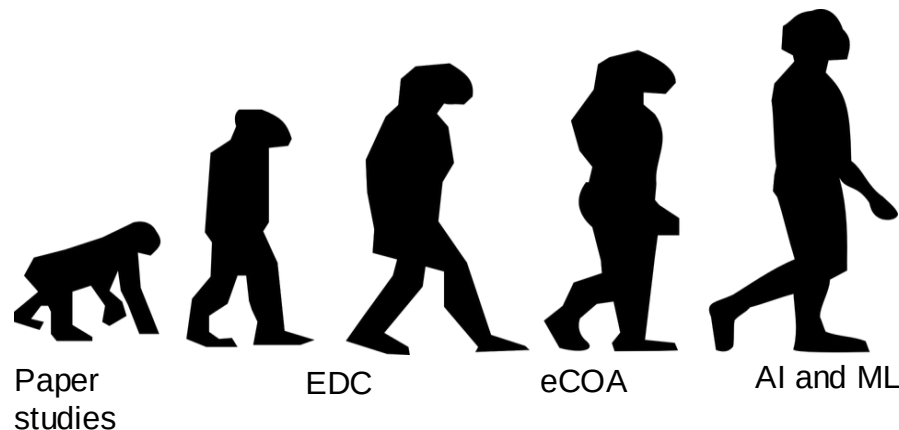
Introduction

- Major changes in the industry have increased the data burden.
- One type of data that poses a particular challenge, is external data.
- It becomes a critical task to transform all these “irregular” data into CDISC compliant data.
- Our main challenges:
 - Diversity of Data
 - Early Phase Time Constraints
 - Management of Vendors and their systems
- Teams will need to adapt, using CDISC standards as our weapons against chaotic data, with regulatory compliance as our final goal.

Main Challenges When Handling External Data

Reyes.elizaldeduarte@syneoshealth.com

Diversity in External Data



Clinical trials now require different types of samples, tests, and assessments, all to meet study endpoints. This has, in consequence, created new types of data with new behaviors, formats, structures, and goals.

The main challenge when facing external data, is that any assessment can be externally sourced

Diversity in External Data – Data Categorization

How it's measured:

- Nominal
- Ordinal
- Interval
- Ratio

How it behaves:

- Discreet
- Continuous

Data attributes:

- Data latency
- Volume
- Velocity at which data is generated
- How structured or unstructured data is
- Veracity
- Accuracy
- Precision
- Completeness
- Consistency

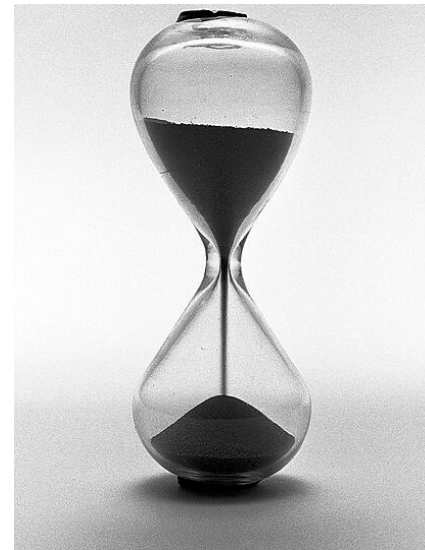
Early Phase Time Constraints

Early Phase (EP) clinical trials are characterized for having short durations.

Diversity of data + EP time constraints = More challenges

- Different data types may result in different structures and formats.
- Little room to resolve discrepancies
- More effort, time, and resources are needed.
- There could be major delays.

All risks will need to be identified and mitigated as soon as possible.



Management of Vendors and their Systems

- Vendors will follow internal standards for data collection and processing and will have a plethora of technologies and systems to their disposal, our combinations become unique.
- Validated systems must be used to generate data, but this does not ensure CDISC compliance.
- Vendor systems are frequently designed and customized to meet the vendor's needs, not the trials or DM's requirements.
- Standard processes include creating Application Programming Interfaces (API) or, data import agreements (DIA) and data transfer agreements (DTA). However, each of these options come with their inherent challenges.

Vendor management becomes a key skill to ensure that external data meets CDISC standards.

Transforming Data To Reach Standardization

Reyes.elizaldeduarte@syneoshealth.com

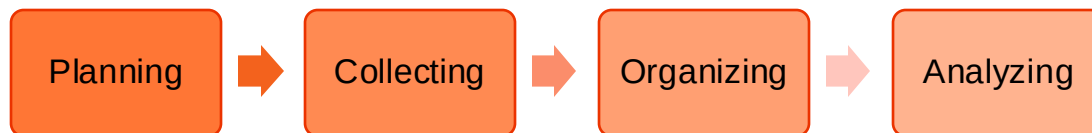
Data Transformation

Can be defined as any change on the data attributes.

It allows data to be effectively used with different systems, processes, or structure.

Data transformation will never mean changing the source value or source meaning of the data. The **traceability** of data needs to be the core concept when transforming data. If traceability is lost, data is no longer compliant.

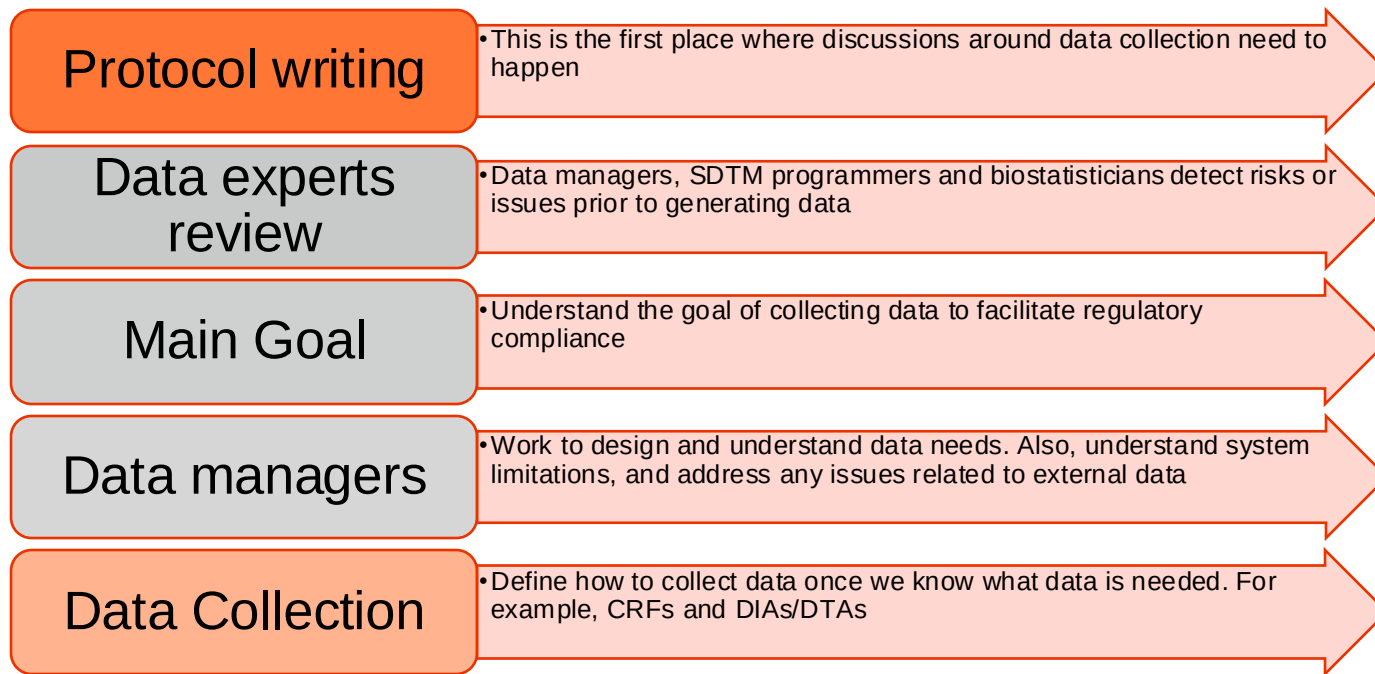
During a clinical trial, four periods can be distinguished:



- Data collection – ‘data shaping’
- Data organization – ‘data conversion’

These steps will be closely connected to each of the CDISC standards, CDASH, SDTM and ADaM. These three standards will be our greatest allies to transform the data.

Plan: What Data And How Is It Needed?



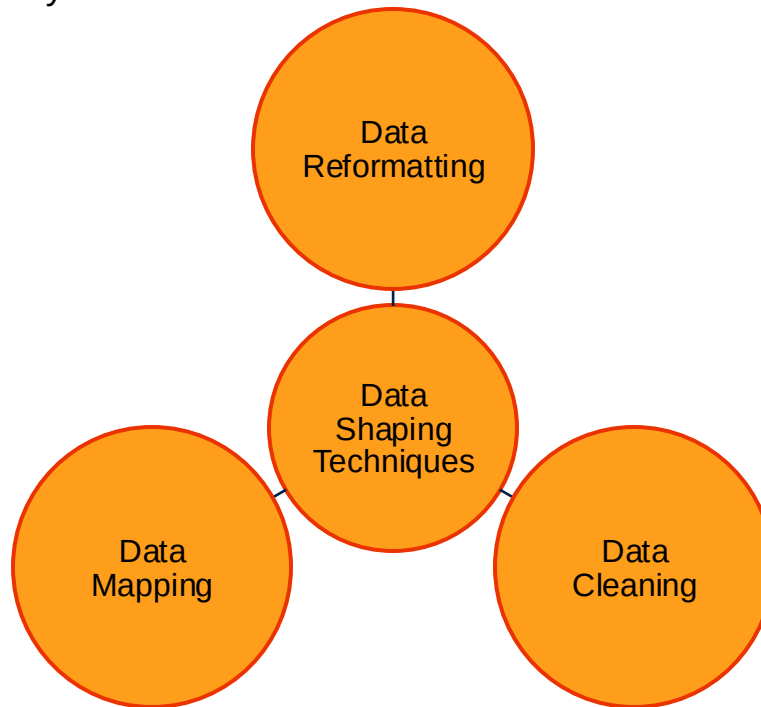
Collect: How Will Data Be Generated And Collected?

Data collection is the main step of our domino effect.

- Define the technical specifications of our data.
- We are still handling raw data. We should never transform data to give it a different meaning.
- Data collected will come in all shapes and forms. CDASH will help us reach standardization.
- We commonly use CDASH during eCRF development. Nevertheless, CDASH standards are just as important for external data.
- When the transfer specifications are built, one of the key points will be to first identify the type of data per the CDASH classes.

Collect: How Will Data Be Generated And Collected?

Unfortunately, not all vendors and systems will be able to meet these requests. This is where our data transformation will come in handy.



Collect: How Will Data Be Generated And Collected?

Combining the flexibility of CDASH to categorize data into their classes and using these data shaping techniques, we can create standards to almost every data type out there.

Proposed solutions:

- DIAs/DTAs Templates
- Data Shaping Specifications and process in place

We tackled and solved all our challenges, ensuring external data is CDISC compliant every time. However, we can still face some more complex challenges that cannot be resolved during this step

Organize: How Will Data Be Presented For Analysis?

Not all challenges can be overcome during data collection. We can face complex issues that require more effort.

- Data organization uses SDTM standards which provide a way to standardize data to facilitate its management and analysis. Following SDTM standards will allow us to warehouse, reuse, share, review and improve data.
- SDTM gives us an additional point of view of the data so additional data transformations are possible.

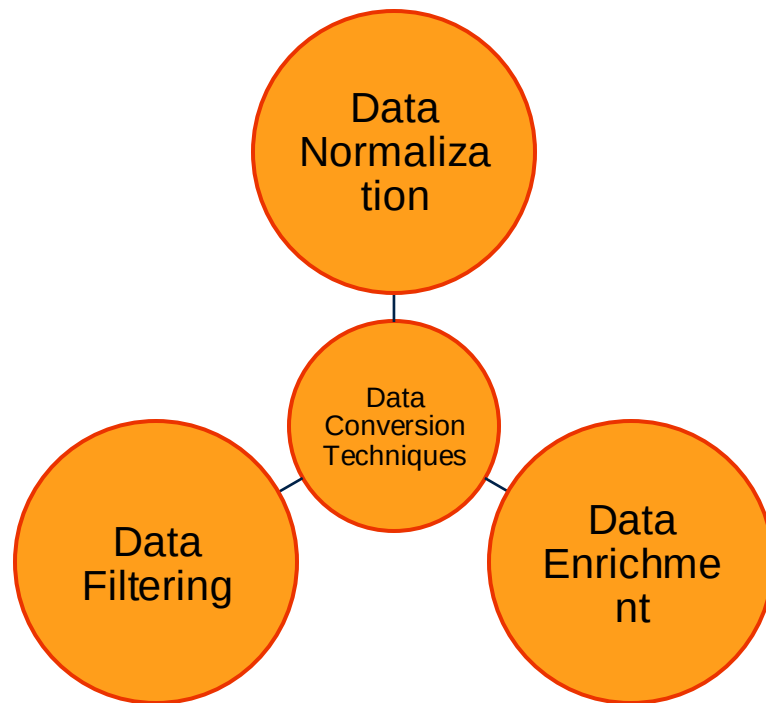
Organize: How Will Data Be Presented For Analysis?

- The data conversions allows compliance when the planning and collection (data shaping) stages were not enough.

Proposed Solutions

- Ensure CDASH compliance prior to organizing
- Data conversions

The data organization stage has given even more solutions to standardize the data and face any challenge, no matter how complex. Data is now ready to move to the last stage, analysis.



— Analyze: Giving Meaning To Data

The last stage on this data cycle is analyzing the data. During this stage there will not be any data transformations as we have already resolved our challenges and made it CDISC compliant.

Data analysis is governed by ADaM standards to perform the needed analysis.

ADaM standards no longer focus on ensuring data readiness, cleanliness, completeness, and accuracy. Therefore, if data is not compliant when shared for analysis, the trial results will not be correct, wasting all the time, effort and resources used.

All stages need to be considered with equal importance, as they are all part of the same cycle. If one stage is not done correctly, our challenges will not be overcome, and the trial will not have the expected results.

Example Scenario

Safety laboratory data is commonly handled using DIAs/DTAs. The vendor will receive the blood samples from the clinical site to perform all analysis. DM already ensured most variables were CDISC compliant, but the vendor has some limitations that need to be resolved.

Data Shaping solutions:

1. Clean all data inconsistent to the protocol by removing it
2. Reformatting the test categories to give the expected names
3. Map the visit names to match the external system with the clinical database.

Data Conversion solution:

1. The Statistical Analysis Plan (SAP) of this hypothetical study, requires US and SI units for the analysis. Normalize data to convert the US units into SI units using established specifications.

Data Analysis:

1. Perform the average calculation of the needed laboratory tests values

Cross Functional Teamwork Is Fundamental

All functional areas (FA) involved have responsibilities and knowledge, but they also have limitations. One FA will not be able to plan, collect, organize and analyze all data, and no FA has the needed expertise on all three standards: CDASH, SDTM and ADaM.

The involvement of critical roles is needed throughout the process. Data managers, SDTM programmers and biostatisticians are particularly needed due to their expertise. Each of these roles needs to be involved during all stages, with varying levels of responsibility, based on the company needs.



Conclusion

- Use CDISC standards
- Defined a clear data strategy within SOPs or processes
- Data traceability
- Teams need to be balanced between stability and flexibility
- How involved each FA or role needs to be during each stage?

Nevertheless, any strategy or path taken should always ensure irregular data is handled correctly, through data transformation, to ensure compliance of regulatory and CDISC standards.

Q&A



Shortening the Distance from Lab to Life[®].



Eduardo Elizalde
Project Data Manager
reyes.elizaldeduarte@syneoshealth.com