

Dealing with multiple data sources: Fundamental aspects of data integration and reconciliation for high data quality



Sonja BANJAC
Data Manager & Analyst
@Kayentis

March 4th 2024



kayentis

kayentis.com

acdm24 
 **Copenhagen**
3rd-5th March



- ✓ What have we learned from the evolving clinical trials landscape?
- ✓ Integration of multiple data sources
 - What we are still learning through multiple data sources integration
- ✓ Case studies: DM practices to support different types of integration
- ✓ The necessary evolution of DM practices for today's clinical trials

What have we learned from the evolving clinical trials landscape

The goals of modern clinical trials

- ✓ More comprehensive understanding
- ✓ Deeper insights and identification of potential trends
- ✓ Accelerating trials
- ✓ Developing effective treatments more quickly

- ✓ Decentralized trials
- ✓ New technology
- ✓ Real-world data
- ✓ Multiple sources



Data diversity and quantity

- ✓ Phase III trials collect three times more data than 10 years ago
- ✓ 20 endpoints on average (Tufts CSDD, 2021)



Site team



Data quality ?



Patients

Integration of data from multiple sources can potentially help with

- ✓ Data accuracy
- ✓ Data consistency
- ✓ Quality of data
- ✓ Compliance improvement
- ✓ Burden reduction
- ✓ Patient and site satisfaction



More data brings more signal, but also **more “noise”**

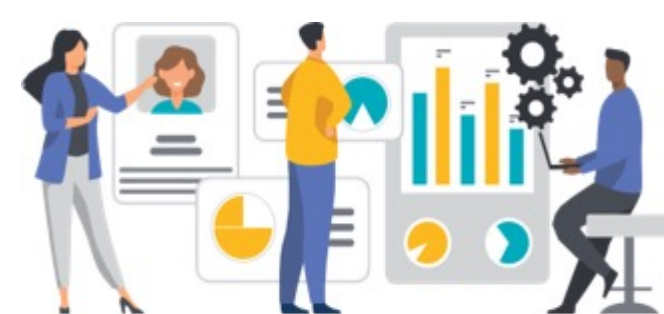
Integration



Data cleaning

- ✓ Identifying errors to ensure consistency
- ✓ Support smooth integration
- ✓ Ensuring data integrity

Data managers (DM)



DM practices to support different types of integration



Current Data reconciliation practices: eCRF and eCOA:

- ✓ Monthly
- ✓ Manually



Shortcomings !

- ✓ Time-consuming
- ✓ Prone to errors
- ✓ Not appropriate for data volume
- ✓ Not in real-time

Solutions

Product solutions

- ✓ Adopting automated reconciliation solution
- + Identifying discrepancies more easily
- + Saving time
- !! Cannot prevent data entry errors

DM solutions

- ✓ **Identifying repeating** issues
- ✓ **Proposing tailored trainings** for these issues (e.g., ad hoc webexes)
- ✓ **Proposing data cleaning** actions improvement
- ✓ **Updating query** texts to increase sites' understanding



Data acquired via

- ✓ Provisioned devices
- ✓ BYOD
- ✓ Web

← **Risks!**

- ✓ Integration of dynamic data in a timely manner
- ✓ Sync issue in case of connectivity issues

Solutions

Product solutions

- ✓ Real-time updates on the platform with the audit trail
- ✓ Accessibility and clear data organization
- ✓ Standard reports available for download to sponsors
- ✓ Automatic inconsistency checks

!! Lack of understanding of the reason behind the discrepancy

DM solutions

- ✓ **Monitoring of automatic checks**
- ✓ **Source check** – analysis of the source of the discrepancy/issue
- ✓ **Instant reaction** - contact with the sites and sponsors
- ✓ **Tailored reports** - based on studies specificities and risks

*Thorough grasp of the study protocol
AND
the technology*

- ✓ Increasingly complex study designs
- ✓ Blinded and unblinded data

← **Risks!**

- ✓ Data security
- ✓ Data integrity

Solutions

Product solutions

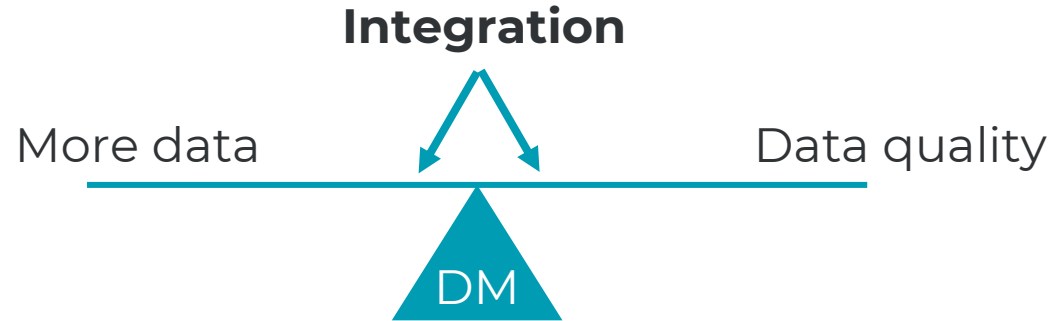
- ✓ Creating different profiles
- ✓ Limiting the access to the platform

DM solutions

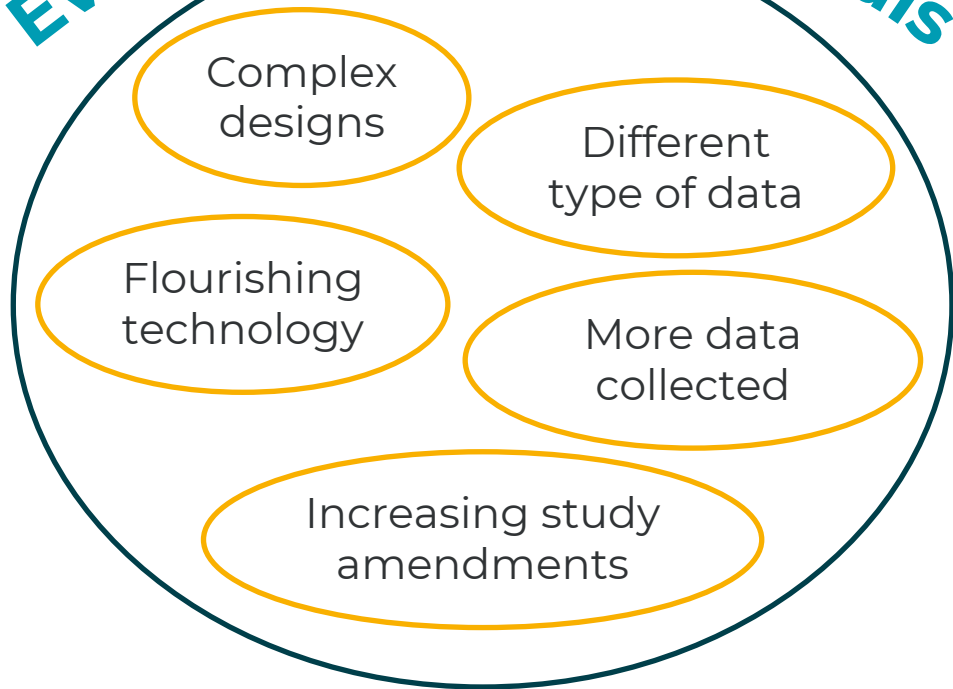
- ✓ **Customized reports** to sponsors
→ monitor blinded assessment, without unblinding the patients
- ✓ **Adapted communication** with sites staff
- ✓ **DM = The “silent link”** between both platforms



The necessary evolution of DM practices for future clinical trials



Evolution of clinical trials



Four main areas for DM skills evolution in new generation clinical trials



#1 Deep understanding of the technology

- ✓ How are the technologies working, communicating and integrating data
- ✓ Various users' perspective
- ✓ Continuing Professional Development

#2 Statistical and informatic literacy

- ✓ Creating solutions to ameliorate data cleaning of multisource databases
- ✓ Combining solutions with used technology
- ✓ Automating repetitive tasks (using ML/IA to uncover patterns)

#3 Strong communication skills

- ✓ Understanding different actors' needs
- ✓ Being able to explain quickly, clearly and plainly the necessary actions
- ✓ Knowing when DM should intervene and when it should be delegated

#4 Inventive and proactive spirit

- ✓ Using their data cleaning experience to propose processes updates, keeping in mind GCP and study protocol
- ✓ Creating reports that help monitor risks
- ✓ Proposing trainings for sites if repetitive discrepancies are identified

THANK YOU



 kayentis