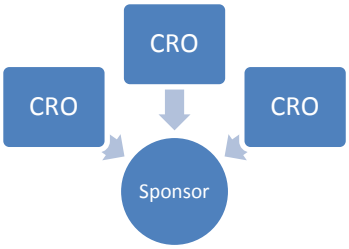
A case study for gaining accelerated access to data and reports through metadata-driven process from EDC to reporting

Challenges

- Data scattered across multiple vendors
- Data in various EDC structures
- Multiple versions of data
- Poor data quality
- Delayed access to data
- Many internal systems to source data

When a major pharmaceutical company was looking to find a solution to harmonize all its data for gaining accelerated access to their data and reports, the company faced a major challenge: data strewn across multiple vendors and in various EDC formats, poor data quality due to conflicting versions, and delayed access due to confusing internal systems and data sources. With a need to handle huge volume of studies in less time, the sponsor wanted a solution that effectively improved the data quality and provided accurate analysis reports for in-time decisions.
Quick access to good quality clinical study reports has always been an important step for clinical teams in any clinical trial. The path from data capture to obtaining statistical reports has been riddled with many obstacles. This case study describes how this pharmaceutical company was able to find a way to accelerate availability of quality data and statistical reports through efficient metadata-driven process for faster decision-making.

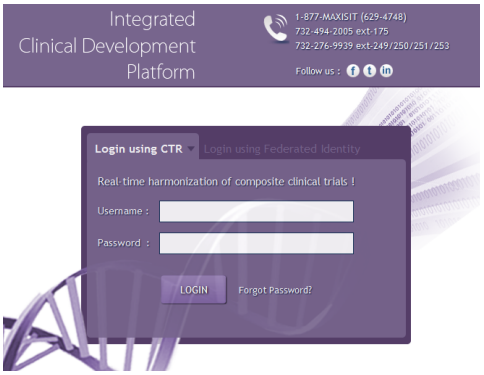
Background



In a typical pharmaceutical landscape, most sponsors work with many CROs and EDC vendors. As a result, the data is captured and stored in various designs and structures. With so much data residing in incongruent formats and silos, it becomes impossible to analyze and interpret data within a unified ontology. Most research organizations and sponsors have realized that having this data in CDISC standard structures help them to easily obtain statistical reports through readily available tools. In an attempt to quickly get access to statistical reports, they pool-in data from various vendors they work with. With disparate interpretations of CDISC rules that might have been applied by their various data providers involved in standardizing data, the quality of data (although in standard format) is not at its best.

eXchange – an integrated data management platform

One of the first steps in building this solution was to be able to lay the groundwork for handling the disparities due to various database designs and structures. MaxisIT's CTRenaissance eXchange provided the platform to handle the level of complexity and volume. Even before the raw datasets from actual studies were standardized to CDISC SDTM format, we identified aspects of similarities and differences between the studies done by various vendors engaged by the sponsor. We noticed that the sponsor had studies captured and stored in many EDC formats. Additionally, for their therapeutic area specific data, the sponsor had their own database structures.

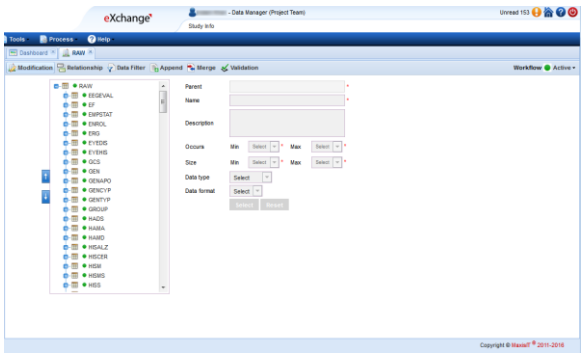


What we did not do

If this were to be handled conventionally, studies that are similar in their EDC database structures would be grouped together, and if SAS programming was used, program codes for previously coded studies would be copied over and reused. Even if metadata were to be leveraged to gain efficiency, Excel based metadata sheets would have been reused to be applied to similar group of studies. For anyone who has used these methods for achieving efficiency, it becomes increasingly cumbersome and difficult to manage various pieces of code and Excel sheets. With huge volume of studies to be standardized for a sponsor this big, the conventional process of reusing code and sheets becomes increasingly unsustainable; the quality of data feeding into the reports poor, and worse, the reports erroneous.

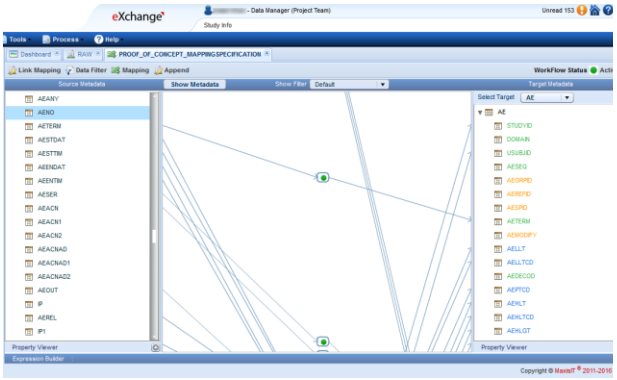
Solution

At this point, having identified all the various database structures, we set out to prepare metadata catalogs for different formats of data. These catalogs are the consolidated list of all datasets and variables that make up a study, and are specific to a particular EDC or sponsor specific data structures. CTRenaissance eXchange provided the ability to import these catalogs into the tool as though they were actual study datasets even before the actual study data was made available. The next step was to create master mapping templates for the sponsor and various EDCs. CTR eXchange provided the ability to apply CDISC SDTM rules to raw study variables and map them to corresponding SDTM variables. The maps were created by simple drag-and-drop tasks requiring no programming.



Master Metadata catalogs in CTRenaissance eXchange

We created master mapping templates for EDCs that the sponsor used, in addition to creating templates for sponsor-specific data structures. We were able to map 70% of raw variables to CDISC SDTM variables, even without handling actual study data. With so much data already harmonized to standard structures at this point, only a minor amount of effort was spent in standardizing the remaining variables based on study specific designs. The master templates were validated and finalized for a specific metadata type. These master templates were then reused to standardize actual study data with the same set of underlying CDISC rules and guidelines applied uniformly throughout. In this process, we eliminated the use of Excel or SAS based programming tools which are cumbersome for metadata modifications.
By leveraging the metadata and reusing mapping templates we were able to reduce the amount of efforts on standardization on any given study by almost 70%. For studies that took months to wait till anybody could actually use study data, the clinical teams were now able to access high-quality data and reports in just over 2 weeks from the availability of raw data.



Reusable Master Metadata mapping templates in CTRenaissance eXchange

Results

The solution provided the capability to the sponsor wherein safety experts were able to quickly access safety reports for numerous studies sourced from many data vendors. It helped them perform in-time deliberative tasks like risk monitoring and safety evaluations for studies going to DSMB/DSUR. Physicians were able to make vital decisions on the course of the trial based on the accumulated data made available to them in formats suitable for their analysis. They were able to use the reports to evaluate safety, study conduct and integrity of the trial. Since the target data was in a homogenous structure, pooling of studies was made possible revealing further important characteristics and insights for group of studies in a particular compound.

Outcomes

- Quick access to reports for safety review
- In-time deliberative tasks during study conduct
- Risk monitoring
- Seamless pooling of study data
- Single repository for all data access and storage
- 70% reduction in overall efforts from data capture to reporting

With this solution the sponsor was able to get quicker access to study reports with an underlying high-quality data stored in a single repository within the CTRenaissance platform; ability to pool homogenized data seamlessly for legacy and ongoing studies, increased efficiency from data capture to reporting, and decreased cost of standardization.
This metadata-driven approach allowed the flow of data from capture to reports at greater speeds, with improved quality, and in the end the sponsor was able to get away with quicker access to statistical and safety reports with real-time decision making.