PharmaSUG 2018 - Paper SS-17

Why organizations need MDR system to manage clinical metadata?

Abhinav Jain, Ephicacy Consulting Group Inc.

ABSTRACT

In the last decade, CDISC standards undoubtedly have transformed all data handling aspects in the sphere of clinical research. Pharmaceutical sponsors and regulatory agencies now have widely recognized standards for data collection, tabulation, analysis, and submission. Sponsors have made a significant investment of time and resources to develop skills in understanding and implementing the standards. As future advancement, clinical research industry is now heading towards utilizing technology to gain efficiency in standards implementation. MDR (MetaData Repository) system is one such tool that helps organizations to manage and centrally deploy the standards.

In this paper, I intend to cover the features offered by an MDR System. How organization, therapeutic area and study-level data-collection metadata can be stored and efficiently managed with MDR system. How an MDR system is relevant to a statistical/clinical programmer. How the SDTM generation processes can be automated using mapping defined in the MDR system with standard CRF forms.

INTRODUCTION

The fundamental purpose of clinical data standards is to streamline the clinical research process with a framework for data collection, data management, data analysis and reporting. CDISC established in 1997 as a standards development organization, partners with industry and regulators to develop data standards to increase patient safety, data quality and transparency.

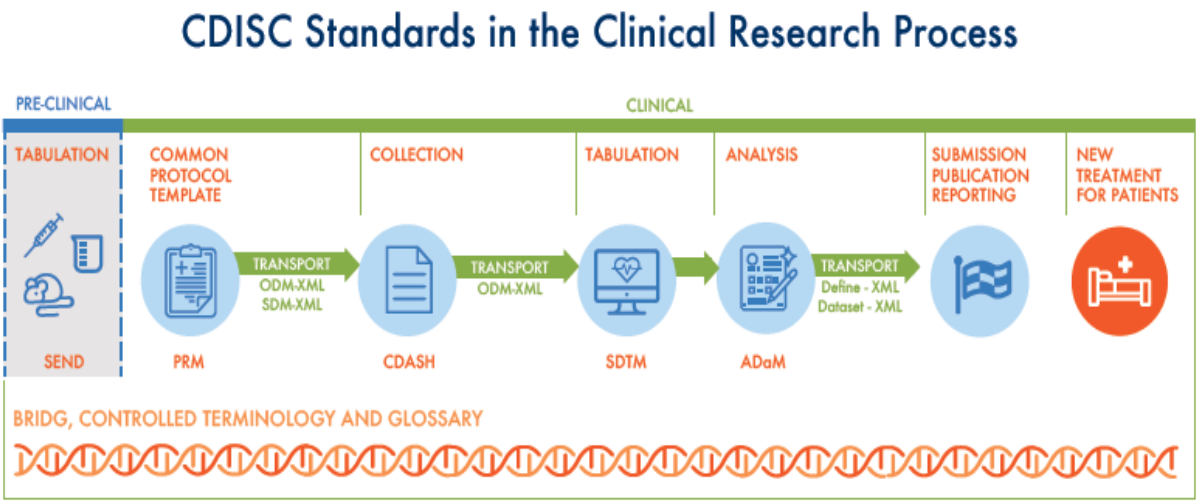

Figure 1: CDISC Landscape

The standards implementation processes with CDISC framework can be harnessed to:

- Ensure traceability of the data from collection to analysis and optimizes the downstream data flow
- Maintains consistency in data storage formats and facilitates data interchange between partners and providers

- Provides uniformity in the creation of analysis parameters thereby reducing the amount of time required for submission and regulatory approval
- Development of automated tools for process automation of data extraction, cleaning, transformation, validation and analysis
- Significant reduction in time and resources for study start-up, analysis and submission.

With evident benefits, still several challenges continue to persist with standards implementation.

- Majority Biopharma companies use spreadsheets and word files to maintain the data definitions. Use of spreadsheet for metadata management results in manual and inefficient processes for content management, change control, versioning and distribution of metadata.
- Standards are implemented in silos and data processing at Collection (EDC), Tabulation (SDTM) and Analysis (ADaM) stages are not linked. Data standards are not implemented with considerations for the successive stage. The leads to the creation of tools and processes that do not gain efficiency from standardizing the predecessor.
- Multiple CRO partners and distributed teams implement standards with different interpretation leading to inconsistent implementation within the organization.
- For large organizations having investigational drugs in various therapeutic areas/disease indications, maintaining global standards and therapeutic area standards is a mammoth task.
- With the release of version upgrades in the Standards, organizations struggle with maintaining multiple versions of the standard.

Implementation of MetaData Repository (MDR) system can conquer the standards implementation and governance practice challenges. MDR system that is integrated, easy to access and with centrally governed definitions have the potential to facilitate end-to-end standards implementation, ensure data compliance and promote automation. A fully functional MDR system will enhance the quality of the submission with the significant reduction in cost and time.

MDR SYSTEM – END-TO-END STANDARDS IMPLEMENTATION

MDR system is centralized repository system and metadata governance platform that can maintain the inventory of metadata definitions for data artifacts originating in various stages in a clinical trial. These metadata definitions include objects like forms, fields, domains, variables, mapping algorithms, controlled terminology, codelists, codes, etc. MDR systems are installed as a local application or a cloud-based application and has interactive GUI interface.

Figures 2 illustrates MDR framework with its collaborative interface facilitating end-to-end standards implementation at collections, transformation and analysis stages.

Organizations constitute a standards team to commission the MDR system. Standards team undertakes the task to establish governance process and working practices around the metadata management through MDR system. Standards team coordinate with clinical operations, biostatistics, data management, programming team and external users to oversee the deployment and application of MDR system in various business processes. This group sponsors the MDR system and works with MDR product companies to select the right MDR product for the organization.

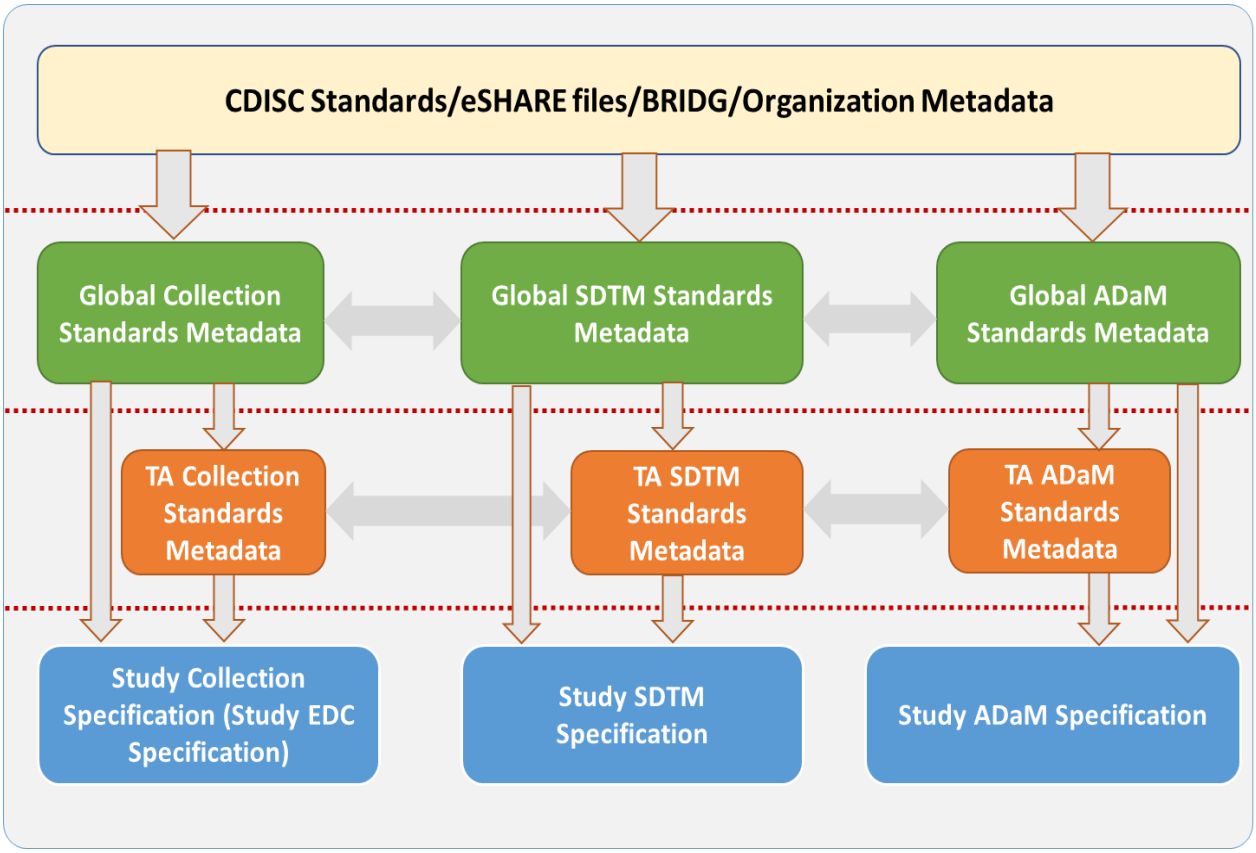

Figure 2 MDR Framework

KEY FEATURES OFFERED BY MDR SYSTEM

- MDR system allows maintenance of standards at hierarchical levels within the organization. This system concurrently manages the Global (Organizational), Therapeutic-Area (TA) and Study-level metadata objects. This system provides the mechanism to define the relationship among metadata objects at various levels. This hierarchical relation can help in analyzing the potential impact of upstream changes in metadata definition on the downstream metadata objects.
- MDR system can store the metadata definition for varied data models: Collection, Transformation, Analysis. Along with EDC, SDTM, ADaM standards, MDR System can store data models for protocol standards, codelists, dictionaries, external data transfer standards (Lab, PK, Imaging, etc.). The system can even manage the multiple versions of a data model like SDTM v3.1.2 and SDTM V3.2.
- MDR system supports the integration of data models to facilitate the end-to-end metadata management. To integrate the data models, the data definitions in predecessor data model are defined with forethought of structural requirements of the successive data model. MDR system stores these relationships in the form of mappings. Once these mappings are established, the system can generate the triggers highlighting the potential impact of the change in upstream model cascading to the downstream data definitions. For instance, ability to analyze the effect of changes in EDC variable attributes updates over SDTM variable derivations.
- In the set-up stage, global metadata objects are established with seeding of industry standards like CDASH, CDSIC Standards/Share metadata, NCI-Codelist. Internal standards and business teams define the data definition rules for global and study-level metadata.

- For ongoing business operations, the standards team manages the MDR system through the governance processes established by the team. Governance process defines the frequency of updates, workflows to be followed for standards updates. Standards team also defines how MDR system will integrate with other organization tools and application.
- Integrated Change control is one of the most sought-after features in MDR system. Through a change control process, a user can log the request for metadata updates. These requests can be about adding/updating/retiring any metadata objects. Change control system constitutes a workflow process that defines the approval process, tracking and handling of change requests. Change control system tracks all changes made to a request from initiation to completion, including what, who, when related details.
- MDR can maintain versions of all metadata objects. For example, different versions of Global SDTM metadata, study ADaM specification etc., when such objects are created or changed. Version control helps to identify which version of the standards is used for a study. MDR allows to visualize the full history of a metadata object.
- MDR system provides access to internal and external users from an organization through an access control mechanism. MDR system has the defined role-types as System Administrator, Metadata Managers, Study Builder. Users are given access based upon their role-type. This feature establishes a control on who can access and update each segment of metadata objects. Since MDR system is a controlled system, external users like CRO partners and other vendors can be given access based on their business requirements.
- MDR system has a built-in audit-trail mechanism to maintain the log of who accessed the system, who made what changes and when the change is made. This feature makes the system traceable and audit compliant.
- To support study programming MDR system creates study artifacts like Study EDC specification, SDTM specifications, ADaM specifications, Define.xml. Artifacts exported out of the MDR system can be used for programming and quality control processes. This eliminates the need for creating and managing programming artifacts in excel spreadsheets or other file formats. Also since the artifacts are available through MDR systems, distributed teams can always access the most recent specifications from a single common source. This eliminates the risk of working on outdated specifications and eradicate the practice of spreadsheet circulation via email.
- MDR system can generate the difference reports that summarize differences in metadata object attributes between specified hierarchies and metadata at various stages (e.g., draft vs. production version). These reports can help in addressing the consistency issues and standards adherence issues.
- MDR system empowers organizations with smooth implementation of business rules consistently across the board. With the algorithm derivations, macro information, programmatic implementation stored in MDR system, these metadata objects are available to all the studies. For example, MDR system when stores definitions of Oncology endpoints, same definitions and derivations for primary and secondary endpoints can be used for all the Oncology studies. The system also supports different variations of the same variable. Among different versions at Global and Project-level derivations, study programmers can choose which variation applies to their study.
- MDR system has inbuilt tools and reporting modules to maintain the quality of metadata, implement internal and between-component consistency checks. These modules generate reports to summarize the deviation and issues with metadata entries. These checks can check the internal dependencies between data models, attribute compliance of metadata objects, the existence of referenced dictionaries items in the global dictionary. For example, if the SDTM algorithm requires a collection variable and if that collection variable is not present.

- MDR system being a storage-house for standards paves a well-defined path for process automation. Implementation of MDR system enables automated creation of study artifacts like EDC specifications, SDTM/ADaM Specifications. With MDR having capability of storing programs, edit check programs there is feature facilitating automatic code generation. With auto-generated programs, programming teams will only have to implement the study specific changes to programs.

MANAGING DATA COLLECTION STANDARDS WITH MDR

In any clinical trial operations, setting up study database and associated activities are always being a process-driven, labor-intensive and time-consuming. Standardized operational (data collection) metadata can significantly boost process adherence and bring efficiency in setting up study database. MDR system manages the data collection standards and automates the creation of study data collection artifacts. At collection-level MDR stores the operational metadata corresponding to EDC framework used by the organization. For example, the operational metadata configuration can be based on CDISC-ODM or Rave architecture.

Case when an organization has Rave based EDC system, MDR system is configured with Rave architectural framework and components. It will have the inventory of metadata objects based on Rave ALS information, including CRFDraft, forms, fields, data dictionaries/entries, unit dictionaries/entries, folders, etc. Along with Rave ALS components, MDR systems will governance attributes around each metadata object. Advanced MDR systems can also include the additional Rave components like edit checks, edit checks derivations.

INITIATING METADATA MANAGEMENT WITH MDR SYSTEM

In the initial stage, Standards team works with the study teams to come up with the standards forms, fields, data dictionary and data dictionary entries definitions. The team then defines the rules and nomenclature about how these metadata objects will be defined and managed. The team establishes the governance process to manage the metadata and reporting requirements to monitor the state of metadata.

Large organizations sponsor numerous studies in diverse compounds and therapeutic areas. In such case, it is not feasible to maintain all the standards at the global-level. Maintaining all standards at global level poses the logistic challenge for standards team and bring inefficiency at the stage of study-level specification creation. To resolve these challenges, it is suggested to maintain TA-level standards along with Global standards. This way standards can be managed efficiently and project teams can also derive maximum efficiency from standards metadata. Though the decision of what should be maintained be at global vs. TA-level should be carefully exercised.

At the Global or TA-level metadata objects are defined with a vision to address the requirements of current and future studies. Seeding CDASH Standards and NCI-codelist is a good starting point to build the organization metadata. Nomenclature rules can be designed to assist in identifying metadata variations at Global, TA and Study-level.

To integrate collection and tabulation data model, it is imperative to pre-define the relation between EDC field and SDTM variable. These relationships can be stored in MDR system through nomenclature or other semantic features. Another critical factor is the utilization of shared codelist/controlled terminology at collection and transformation level. Codelist management through an MDR system is a vital component of a successful MDR implementation.

STUDY SPECIFICATION WITH MDR SYSTEM

At the study-level, study DM works in conjunction with Biostat, Programming, Clinical Operations team to define the data collections requirements for the study. As an outcome, the team identifies the CRF components forms, fields and associated field-values corresponding to study requirements. Next step, using the MDR system study DM builds the data collection specification by consuming Global and TA-level metadata objects like standards forms and other associated components. During the development stages, if any standard metadata objects need an updates study teams can request those changes via MDR Change Management system.

At times there are situations when all the study required data collection elements are not standardized and are not defined in Global and TA-level collection metadata at the time of study creation. In that case, study team can define those components at the study-level. Using the export feature MDR system generates the collection specification in a format that can be directly consumed by the EDC system. Since the MDR system supports the metadata management at study-level, final study collection specifications can reside in the MDR system. For efficiently administration of standards implementation, study team should diligently use the standard forms and fields unless available standards do not meet the requirements of the study.

Case when the changes to the specifications made outside the MDR system, using the import feature specification can be imported into the MDR system. As the study progresses subsequent updates to specifications can be managed within the MDR system utilizing version control features.

Using the MDR system comparison features, the system supports the generation of difference report comparing the study collection specification with global, TA-level metadata or between two versions of study collection specifications.

Inbuilt tools in MDR system can check the standards compliance of study specifications and report deviations. Various reports can be generated to report the deviations. For example, all mandatory fields are included in the form, all conditional fields meet the conditional requirements, data dictionary terms have all the terms included from global codelist, nomenclature rules are followed.

Through the governance process, custom changes made at study-level are adjudicated to be added to Global or TA-level metadata. With this cyclical process, metadata can be enriched and subsequent studies can benefit from the customization that was done for previous studies. This progressive elaboration of metadata definitions leads to the continuous reduction in time requirements for study set-up.

MANAGING SDTM STANDARDS WITH MDR

Post collection next stage of data processing is data tabulation. CDISC SDTM standards are now a globally accepted standard for clinical data tabulation. Conventionally SDTM specifications documenting mapping relation between EDC fields and SDTM variables are developed in excel spreadsheets. Managing specifications developed in excel spreadsheets are carried over to futures studies and custom changes are made with every study. With this approach mapping defined and business knowledge acquired at every stage gets vanished. Variations in the database design of every study also hamper reusability.

Contrary to the conventional approach, MDR system supports integrated management of SDTM metadata. Collection metadata in standard CRFs is defined with a perspective mapping to target SDTM domain. SDTM metadata is defined with the form and field-level information that will be the used to derive that SDTM variable. This mapping information will be like which CRF form will be mapped to which domain, which EDC field will be mapped to which SDTM variable, list of all EDC variables will be used for an SDTM variable derivation.

Once established these mappings relationships are available to be used by all future studies. These predefined, reusable mapping removes complexity and led to consistency among studies.

MDR system is configured to support all components of Define.xml that includes metadata at: domain-level, variable-level, whereclause-level, valuelevel-level, method-level, comment-level and cterm-level, dictionary-level. Standards team works with the study teams to define the metadata definition for each SDTM component at global and TA-level.

At the study-level, target SDTM domains can be identified by linking study EDC specifications with EDC-SDTM mappings. Similarly, by using study CRF field values and mappings associated variable-level and value-level metadata can be populated. SDTM Codelist can be developed by using the study DataDictionary and DataDictionaryEntries. Using this approach MDR system can generate reasonably partially complete draft SDTM specifications.

Study teams can finalize updates to the draft specification within the MDR system through the GUI interface or by exporting the draft specification outside the MDR system. Using the import capabilities, version control and difference reporting tools subsequent versions of study SDTM specification can be maintained in the MDR system. With workflow management system, creation, review and approval of specification can happen within the MDR system. With built-in checks, MDR system generates deviation reports reporting inconsistency between SDTM specification components, the difference between study metadata with global metadata, codelist issues, nonconformity to the standards.

With built-in code generation capability MDR system can generate the partially complete SAS programs. Programmers can make the study specific updates and complete the program with customized study specific components.

MDR system can interface with Pinnacle21 to generate the submission deliverable including Define.xml and submission Pinnacle21 Validation of SDTM deliverables.

Implementing SDTM programming assisted with MDR system can significantly bring down the SDTM development time. Broadly SDTM development timelines can be brought down up to 40-50% with the implementation of MDR system. This efficiency can be maximized with the maturity of metadata definitions.

CONCLUSION

Implementation of clinical data processes with metadata-driven MDR system will help organizations to significantly reduce cost and condense development timelines. MDR implementation promotes reusability, process automation and end-to-end data traceability. Organizations can achieve excellence in data handling capabilities as well process standardization, operational efficiency by commissioning MDR system.

REFERENCES

CDISC Brochure. (https://www.cdisc.org/resources/cdisc-brochure)

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Abhinav Jain
Ephicacy Consulting Group
(860)-990-7028
abhinav.jain@ephicacy.com; abhinavjn@yahoo.co.in