VIEWPOINT

BEST PRACTICES FOR CTMS INTEGRATION PROJECTS

Clinical Trial Management Systems (CTMS) have become an essential tool for clinical research organizations and pharma companies for streamlining and managing complex clinical trials. As the demand for clinical trials grows, the need for efficient and effective integration of CTMS has become increasingly important. Its also pertinent to note that other downstream systems in clinical, regulatory and safety domains consume CTMS data for executing business critical processes e.g. regulatory submissions and safety reporting.

This paper makes an attempt to highlight some of the best practices for CTMS integration, focusing specifically on data migration projects which are performed as a part of mergers and acquisitions. By following these best practices, organizations can execute the projects with reduced risks, faster turnaround and adherence to regulatory and compliance requirements.

Infosys®
Navigate your next

The Problem Statement

Throughout the lifecycle of a CTMS integration project, effective management of CTMS data across multiple sources is critical due to its volume and complexity. The data might need to be sourced from multiple clinical research organizations (CROs), each having its own set of systems and data standards. Data sources include electronic data capture (EDC) systems, source CTMS and Trial master file (TMF) documents such as 1572 forms, delegation logs, investigator profiles etc.

The Fragmentation of data leads to:

- Data inconsistencies (gaps) and errors.
- Inefficient data processing and analysis leading to schedule slippages and cost overruns.

Infosys POV on the Possible Solution/s

A CTMS integration comprises of the following phases.

Discovery

The primary objective is to identify the stakeholders, processes and data sources. The project kickoff should be performed with all concerned stakeholders followed by a few sessions with each stakeholder. The key outcomes include roles and responsibilities of all stakeholders, list of standard operating procedures to be followed, documentation of data standards of both source and target systems, listing of data sources and major program milestones.

In addition, high level estimates of data volume should be determined which should help in assessing the effort and duration needed to perform the CTMS integration. Some of the information about clinical studies if not readily available can also be collected from external sources e.g. https://clinicaltrials.gov/.

Illustration: Data Volume Estimates

- Number of studies to be migrated by their phase and completion status
- Number of sites to be migrated
- Number of investigators and other clinical staff
- Number of organizations

Scoping and Planning

In this phase, the scope of CTMS integration must be finalized and signed off. For large scale programs, it might be prudent to split the integration into multiple phases based on priorities defined by business stakeholders. Scoping and planning can be performed in parallel by conducting workshop(s) with key stakeholders.

The potential challenges, key dependencies and risks (process and technical) must be identified. Possible solutions and mitigation strategies should be worked out and finalized and subsequently incorporated in the project plan.

Illustration

- The safety data must go live on 1-Sep. To achieve this milestone, CTMS must be go live atleast 2 weeks before 1-Sep because safety systems consume CTMS data. This dependency must be incorporated in the project plan.

- A particular study has been recently completed. Clinical Study Report (CSR) submission is due in next 2 months. The study data must be migrated to CTMS on priority. This task must be explicitly planned and tracked.

- The target CTMS sources the drug and IND details from the regulatory information management system (RIMS). In such case, migration of studies can be completed after drug and IND details are made available in RIMS. This dependency must be incorporated in the project plan.

- The patient enrollment attributes might have different definitions between the source and target e.g. First Subject First Visit (FSFV) date in the source is considered the date on which the first patient was enrolled. However, in the target system, FSFV date is the date on which the first informed consent form was signed. In such case, a business decision must be made on how the FSFV dates should be mapped between the source and target systems and business rules and data transformation rules(if any) should be documented accordingly.

- The organization and investigator listings for China and Japan sites are available in Chinese and Japanese respectively. They must be translated to English as the target CTMS does not support Chinese and Japanese languages. Sufficient time must be planned for translation activities.

A robust risk management and governance framework is key to the success of the project. Since multiple stakeholders are involved including those from clinical study teams, upstream and downstream systems and external parties e.g. contract research organizations (CROs), all risks must be centrally managed and maintained in Risks, Assumptions, Issues and Decisions (RAID) log. They must be assigned and monitored by the project management office (PMO). Escalation procedures and key contacts must be well established and agreed upon with all stakeholders.

Figure 1: Risk Management Framework

Risk/Issue Analysis & Classification

- Impact and probability evaluated.
- Priority is assigned.
- Central RAID log is updated based on relevance.

Risk	Impact	Probability	Mitigation Approach
Delays in System Access	High	Low	• Identify key stakeholders and root cause behind the delays. • Address the root cause and escalate as needed.
System outages and interruptions	High	Medium	• Develop a comprehensive plan to address service interruptions. • Provide adequate trainings to end users to minimize the impact.
Data Loss/ Data Breach	High	Low	• Follow strict security protocols. • Monitor and Audit user access management. • Conduct risk assessments and vulnerability testing.
SME Attrition	Medium	Medium	• Plan for early transition. • Identify key personnel and plan for backup resouce.

Risk/Issue Mitigation & Resolution

- Team works on risk mitigation.
- Unresolved risk are escalated.
- Once mitigated, it is communicated to relevant stakeholders.

Escalation Process

PMO
- Review inter-workstream issues.
- Facilitate decision making in order to resolve the issues.
- Issue guidelines to prevent issues from occurring in future.

Program Manager
- Resolve intra-workstream / inter-workstream issues.
- Document inter-team issues in weekly status meetings.
- Escalate unresolved issues.

Workstream Mead
- Resolve intra-workstream issues.
- Escalate unresolved functional issues.

Risk/Issue Monitoring & Reporting

- All the risks are monitored and reported.
- RAID log is tracked to closure.

Risk Assessment - Program

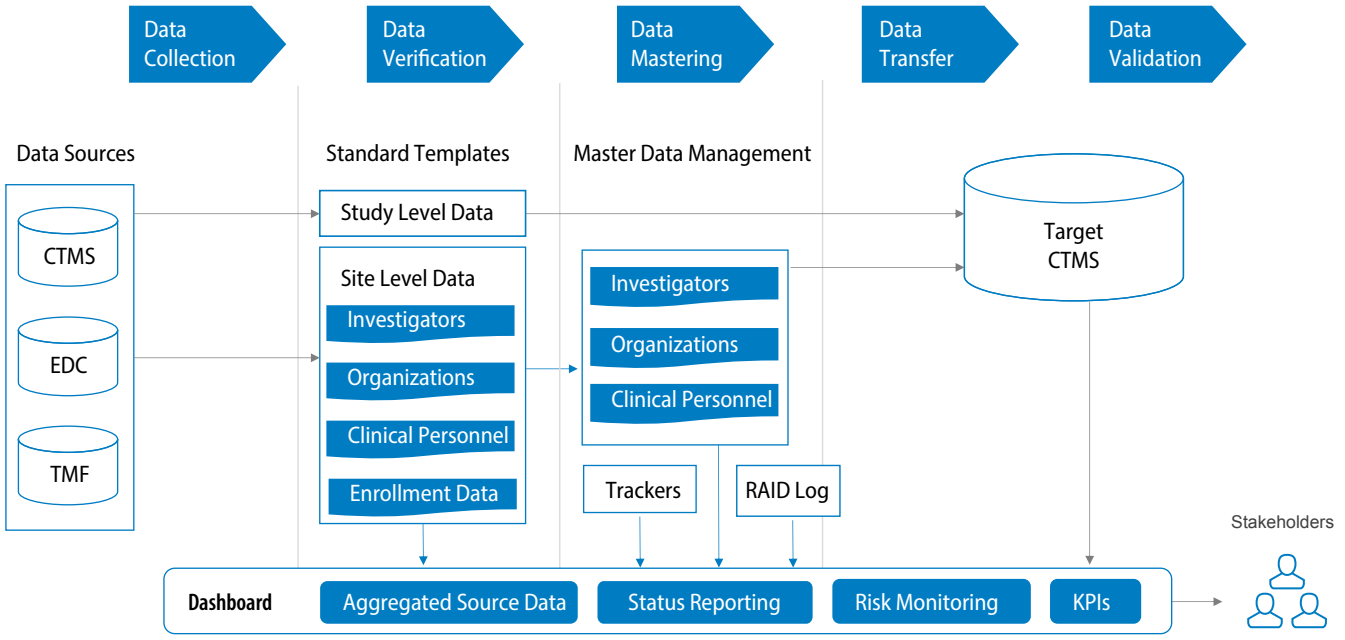
ID	Type	Risk Description	Last Reviewed	Expected Completion Date	Actual Completion Date	Impact	Probability	Risk Rating	Mitigation Plan
1	Data Quality	Email ids are not availabe for 60% of investigators	03/03/2023	15/03/2023	12/03/2023	Medium	Likely	High	Use dummy email ids for inactive investigators
2	Data Quality	Discrepencies in investigator status between listing and 1572 documents	07/05/2023	30/06/2023		Medium	Possibily	Moderate	Get clarifications from CROs and request for updated source documents
3	System Outage	Frequent outages is impacting data entry activities	06/04/2023	30/06/2023		Major	Likely	Critical	Bug fix patch upgrade to be completed on priority
4	System Dependency	INDs not available for 2 ongoing studies (A3453212, B4532222)	15/06/2023	07/07/2023		Minor	Unlikely	Low	Confirm the INDs with regulatory lead

Impact

Probability	1-Minor	2-Medium	3-Serious	4-Major	5-Catastropic
A-Almost Certain	Moderate	High	Critical	Critical	Critical
B-Likely	Moderate	High	High	Critical	Critical
C-Possibly	Low	Moderate	High	Critical	Critical
D-Unlikely	Low	Low	Moderate	High	Critical
E-Rare	Low	Low	Moderate	High	High

The diagram below depicts the data-flow process which is described in the sections below.

Figure 2: Data Flow Process



Data Collection from Source System(s)

Data extracts including protocol data and site data should be obtained from the source CTMS. In case, a single source CTMS is unavailable, data will need to be collected from multiple sources, some of which are listed below:

- 1572 forms
- Organization listings
- Investigator listings
- Delegation Logs
- ICH-GCP statements
- List of clinical personnel (e.g. coordinators, pharmacists and monitors)

One of the key challenges is to extract the data from the sources accurately and ensure that it adheres to the data standards and quality checks enforced in the target CTMS.

We recommend maintaining standard data collection templates based on SoPs and CTMS data standards and perform verification checks once the data is obtained. The templates must ensure that all entities and data fields are configured as per the target CTMS data standards. Also, all requirements for data transformations must be well documented including the business rules and examples.

Data from various sources, such as clinical trial management systems, electronic data capture systems and legacy databases, are extracted and stored in the respective data collection templates after enforcing the necessary data transformations. Generative AI solutions can prove to be very efficient in extracting data from unstructured data sources e.g. 1572 files and delegation logs.

Illustration:

- Let's consider the investigator data set which can have only two roles in the target CTMS namely, "Principal Investigator" and "Sub Investigator". A quick scan of the source data indicated that it contains the values "Principal Investigator" and "Investigator". Hence, there is a need to document and enforce the data transformation rule for converting the role "Investigator" to "Sub Investigator".

- The site status values and definitions can be different between the source CTMS and target CTMS. In such cases, multiple parameters might need to be considered to perform the data transformations e.g. site activation date, FSFV, LSLV, subjects enrolled etc.

Data Verification

Verification checks help in early identification of data quality issues which could be related to date formats, data types, missing values and field lengths. In addition, business rules which are often enforced in CTMS can also be executed to find any data integrity issues.

Illustration:

Let's consider that the email address for investigator records is a mandatory attribute and each investigator must have a unique email id. In cases, email id related issues are found, queries must be raised with the partner company and valid email ids must be obtained. There could also be cases wherein email ids are unavailable for historical records. It is essential to establish a clear protocol for dealing with such scenarios including the option to generate a dummy email address if necessary. The business stakeholders must be consulted to determine the best approach and decision making.

Let's also consider some examples of business rules which must be met by enrollment data e.g.

- Number of patients who completed the study must be lesser or equal to number of patients screened.
- Site activation date must be earlier than patient enrollment date.
- First Subject First Visit (FSFV) must be earlier or equal to Last Subject Last Visit (LSLV) date.

Data Mastering Process

The data mastering process is a critical step which includes consolidation of data from multiple sources and finalizing the records before they can be migrated to the target CTMS. The goal of the data mastering process is to ensure that the data is accurate, complete, consistent and unambiguous.

To achieve this, we employ a multi-step approach that involves the following activities:

- Identify keys for master data entities e.g. investigators and organizations which can be used for identifying potential duplicate records and for building relationships across data sets.
- Perform a comprehensive check against the target CTMS to determine which records are already available. Such records must be re-used instead of creating duplicate records during data transfer. There might be a need to update existing records based on the latest data.
- New records should be created for which no matches were found in the target CTMS.

Illustration:

It was observed that there were inconsistencies in the investigator's last name across different study records which were collected from multiple sources. To address this issue, data mastering techniques were deployed, utilizing unique identifiers such as email address and "study conducted at" organization and geographical attributes (city, state and country) to identify such records. Clarifications were sought from the partner company and the records were updated and finalized accordingly.

Data Transfer to Target System

This process involves the transfer of data from the source system(s) to the target system, ensuring that all relevant data is accurately and efficiently migrated.

The preferred approach for data transfer is to automate the process by leveraging standard APIs which are offered by the target CTMS or developing a custom tool as per the data transfer requirements. Rollback mechanisms must be in place and tested before automated data transfer is initiated. In many projects, data is transferred in batches. It is advisable to maintain data snapshots and logs of the incremental data transfers.

However, in some cases, automation might not be feasible or preferred by the client and data transfer will need to be performed manually using the CTMS user interface.

Comprehensive user training must be conducted and work instructions for entering data must be well established.

Any errors which are encountered during data transfer must be duly documented and reported. Root cause analysis and corrective actions must be implemented to address such issues and prevent them from occurring wherever possible.

Data Validation

Data validation is a crucial step which ensures that the transferred data is accurate, complete and consistent across both the source and target systems. This process involves verifying the migrated data against the original source data and identifying any discrepancies or missing information. The work instructions, data quality checklist and issues tracker must be maintained throughout the data validation phase.

This activity can be largely automated and performed in parallel with data transfer activities. In cases, data from unstructured sources need to be referenced, a random sample of data can be identified for data quality checks which can be manually performed.

Dashboard

Dashboard is an essential tool which can be used to meet the following objectives:

Project Tracking: Provides a birds-eye view of the progress made and also enable drilling down the metrics at the study and site level.

Risk Management: Highlights the high impact risks and their status which are crucial for decision making.

Enhanced Collaboration: Stakeholders can access near real-time CTMS data which is aggregated from multiple sources. This approach helps them to proactively plan downstream activities e.g. trainings and provide valuable insights including early detection of data quality issues and risks. It also helps in cutting down communication efforts for data sharing requests.

Quantitative Project Management: It is important to measure the process performance to implement continuous improvements. This can be achieved by defining key performance indicators (KPIs) and monitoring them as illustrated below.

Data Insights: Highlight important data characteristics and patterns e.g. average number of investigators per site and distribution of sites based on site status and country.

Examples of KPIs

- **Data Accuracy:** Percentage of migrated data that meets quality standards (e.g., 98% accuracy).
- **Data Volume:** Total volume of migrated data (e.g., number of records, GBs).
- **Progress Indicator:** Completion of data transfer and data validation of data sets for each site and rolling it up to study level and project level.
- **QC Error Rate:** Monitor the number of errors encountered during data migration, with a target of less than 0.5% error rate.
- **Top Causes for QC Errors:** Top categories under which 80% of the QC errors fall.

Additional Recommendations

- Conduct regular training sessions with stakeholders on the new CTMS system.
- Establish a post-go-live review process to identify areas for improvement.
- Develop a disaster recovery plan in case of system failures or errors.

Implementing these best practices, one can ensure seamless CTMS integration with high transparency and productivity for tracking project progress during the migration.

Authors:

Ratul Sarkar

(Principal Consultant)

Deepika Sawant

(Senior Associate Consultant)

For more information, contact askus@infosys.com

Infosys®
Navigate your next