CHAPTER 5

FUTURE PROSPECTS

With fast-pace adoption of the technological innovations, CDM role is all set to evolve into a key cross-functional knowledge driven domain connoisseur. The spectrum of skill sets which CDM team can support is expected to get broaden and gradually this will take a step forward to cater to the additional responsibilities like safety signal identification, social listening, RBM, big data analysis etc. Considering this paradigm shift, subsequent sections briefly defines the outlook of pharma fraternity, as expected from CDM [106].

5.1 Metrics for CDM as Next Practice

Data quality, productivity (an activity cycle-time and cost reduction) are key considerations in drug development process and to maintain competitive edge to achieve the organization’s goal. While implementation of clinical data management (CDM) activities, in a pilot study it emerged that multivariate metrics if created based on the application of techniques of simple predictive/exploratory analysis will be of immense assistance to monitor performance of critical procedural steps of CDM. Quality process metrics captures operational performance in terms of how something is being done relative to the known standards or practices to be established that may come from either internal or external sources. Metrics will not only prove useful for the smooth and consistent CDM performance with expected data quality standards, but also to extricate, through timely gap analysis and by providing lodgings for standardization of CDM practices, the potential regulatory risk such as measuring noncompliance to a protocol and SOPs.

Following are some of the example of key metrics dashboards depicting major performance indicators as pilot used cases; recreated based on the replica of original version using dummy data [16].
Figure 5.1 depicts the comparative progress in the process of finalization of CRF for different studies over a period of 6 weeks.

For study-7 at week-3 only six pages could be finalized, as against the study-4 where twelve CRF pages have been finalized.

![Figure 5.1: Advancement in the process of CRF Finalization](image)

Figure 5.2, shows the use of New vs. Existing questions in the Glib for different studies in the respective study databases; thereby exploring the possibility of redundant questions, if any.

![Figure 5.2: Glib Questions, New Vs. Existing](image)

Of note, in study-7, 900 existing questions have been used as compared to 63 new questions created in the database. Also, the maximum number of new questions has been created for study-1.
Figure 5.3, gives the comparison of data entry status in studies running simultaneously.

![Figure 5.3: Depicting Data Entry Status](image)

For study-2, 100 CRFs has the status of 1st Pass Pending, 400 have the status of 1st Pass Complete, 140 has the status of 2nd Pass Pending and 260 has the status of 2nd Pass Complete.

Figure 5.4, gives the comparison of error rates in studies running simultaneously. It is clear from the bar graph that the maximum error rate was observed for study 6 in QC report of Validation Procedures.

![Figure 5.4: Error Rate Comparison of different studies: QC Procedures validation report](image)

Figure 5.5, analyze the error rate for study 6 for various QC reports. The minimum value for error rate was observed in QC reports of ‘Data Extraction Views’ and ‘100% QC of Critical Data Points’.
Thus metrics can promote the transparency of data-quality assessment reporting, and useful to achieve consistence performance and improved efficiency through timely identification of outliers, primarily based on-procedural gaps, upsurge in error rate within the methodology, systematic errors, metadata problems, missing data and incomplete documentation [16].

5.2 : Process Automation, CDM Digitalization and eCRFs

Automation of processes that capture the work flow of CDM can promote the transparency of data-quality by minimizing need for human intervention and manual review. This would enable streamlined communication with efficient accountability. Cross-functional stakeholders can independently ensure adherence to SOPs, study protocol, and GCP. Any deviations, if any, can be rectified through timely intervention with comprehensive insights of the work status along with clearly defined approval hierarchy [107], [108].

CDMS, CTMS (Clinical Trial Management System), PVG and Statistical Analysis databases should have seamless integration in conjunction with the technological advancements, for example automated e-mail and SMS alerts prompting the designated task owner to meet the TAT, mobile apps can help to monitor the work progress for example, discrepancy resolution status, aging of queries etc. Though the initial investment for automation may be high but return on investments are huge as work efficiency is expected to increase by many folds thereby promising decline in cycle time, decreased operations cost, improved compliance and better quality.

Use of eCRF is gradually increasing in clinical trials as compared with traditional pCRFs, depending upon the study specific requirements [18].
Figure 5.6 depicts the diagrammatic representation of proposed process of automation in CDM along with the associated benefits for each of the individual CDM task.

<table>
<thead>
<tr>
<th>Study Startup Phase</th>
<th>CDM Activities</th>
<th>Target Outcome</th>
<th>Proposed Future Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMP</td>
<td>-Decline in Time to create DMP</td>
<td>Automation Creation of DMP</td>
<td></td>
</tr>
<tr>
<td>CRF Designing</td>
<td>-Decline in errors in CRF designing</td>
<td>Automation CRF process</td>
<td></td>
</tr>
<tr>
<td>Design Of Database &amp; Edit Specification</td>
<td>-Use of Standard pages &amp; codes -Reduced Time for UAT</td>
<td>Standard Checks to be Automated</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study Conduct Phase</th>
<th>CDM Activities</th>
<th>Target Outcome</th>
<th>Proposed Future Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double Data Entry</td>
<td>-Decline in entry errors -Less time for Data Entry</td>
<td>Scanning of paper CRF or use of eCRF depending upon the study need</td>
<td></td>
</tr>
<tr>
<td>Query Discrepancy Management</td>
<td>-Decline in secondary queries -DCF Tracking log development</td>
<td>By timely forecast of errors through metrics -Automatic generation/updation of DCF tracking logs</td>
<td></td>
</tr>
<tr>
<td>Coding with MedDRA</td>
<td>-Consistent coding</td>
<td>-Automation of coding</td>
<td></td>
</tr>
<tr>
<td>Lab Data</td>
<td>-Error free Lab data</td>
<td>Batch Loading of Lab Data, directly from Labs</td>
<td></td>
</tr>
<tr>
<td>PVG SAE Recon</td>
<td>-Error free fast SAE Reconcile</td>
<td>Seamless Integration of CDM &amp; PVG database, independent of the platform/vendor</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Closeout (Report) Phase</th>
<th>CDM Activities</th>
<th>CDM Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Export</td>
<td>-Creation of views with no duplication/omission and CDISC format</td>
<td>Automation of process -end users can extract data without depending on CDM. -Automatic generation of Output in a desired format</td>
</tr>
</tbody>
</table>
**Automation: Study Start-up**

- It is recommended use of eCRF for vaccine trials, at least for the centers with support infrastructure. This will allow decrease processing time at the study conduct phase though initial start-up time would be comparatively more. Of note, eCRFs and pCRFs are used in studies with different subject numbers and study sites, due to the differences in overall associated risk; eCRF are more advantageous in large and low-risk studies [18].

- With the use of standard templates, DMP can be generated by the system. This can be done by entering the values for study specific variables.

- Final version of the Annotated CRF can be generated from the system after database designing.

- Similarly process automation can be deployed for DVP documentation. Also, all applicable validations programs/views can be simultaneously copied along with the global standard pages; thus separate QC procedures would not be required for database setup and validation programming. Further, this task can be combined in UAT. Thereby all designated system users can test not only database design and validation checks but also the user rights associated with various study roles.

- Parallel to all the activates, desirable documentation for MDMF for common pages for the studies, can be created automatically by entering the values of the placeholders, for example:
  - SEC document
  - relevant QCs reports for each of the above listed task
  - Handling of Lab Data (Non –CRF Data)
  - Data Coding Guidelines
  - QC Reports

**Automation: Study Conduct**

- For eCRFs, double data entry and reconciliation of data entry mismatch, will not be needed as real time data would be directly entered into the database. Scanned CRFs of paper studies scan be uploaded in the system.

- ‘CRF Data Tracking Log’ would not be required as tracking could be done through the system by flagging out missing/blank pages.
- For data entry at the site DEG document will not be needed for eCRF. Instead CCG (CRF Completion Guidelines) would be associated with each field as the help text.
- Errors associated with single field will be popped-up at the time of data entry and therefore can be resolved at the site on real time basis, for example, wrong data type, incorrect length, mandatory fields left blank, incorrect outliers etc.
- Batch job for validations will trigger discrepancies for multivariate checks and same could be handled at the site. Thus there would not be any need for DCFs, query resolutions with dated signatures and DCF tracking logs.
- Discrepancy Management process automation would allow for selective review of queries for CDM team, for example associated with programming errors, response check etc.
- Data entry page status and discrepancy status can help to judge the progress on real time bases by all cross-functional teams, thereby avoiding the need for creating progress reports for the tasks by CDM.
- Lab Data and other Non-CRF Data can be entered directly through batch load updation. Thus avoiding possibilities of errors. Provisions should be there to upload reports for example, ECG XML, CT or MRI Scan.

**Automation: Study Closeout**
- All of the data extract views will get copied along with standard eCRF pages, thus there would be the minimum need to recreate views.
- Data Coding will be done automatically with the use of Thesaurus Management System.
- SAE reconciliation in CDMS with PVG database will be done automatically with seamless integration of the two systems.
- Before lock, most of the information in ‘database lock checks list’ and DHR can get automatically filed based on the study progress data from the system.
5.3 Patient Enrollment and Study Site Selection

In clinical trials, subject enrollment is a big bottle neck for study progress. The trial may get delayed if the correct sample size or appropriate population for the study is not achieved. It is extremely important to identify correct investigator’s site so that the desired subject enrollment rate could be achieved.

Company’s internal historic data from CDMS can serve as a useful resource for extracting information about the site’s previous performance in patient recruitment. Application of advanced analytics for pattern analysis can be used to deduce information regarding patient recruitment that can be extrapolated to predict sites for future trials [109].

Figure 5.7 depicts diagrammatically information about the sources of information (CDM databases), to achieve desired subject enrollment number for the trial

![Diagram](image)

**Figure 5.7**: CDM databases: Information sources for site selection and patient enrolment [109]

Data managers can provide comparative matrices for similar studies categorizing information in leading (signals future events), lagging (that follows an event) and coincident (occur at around same time) variables based on historic or real time performance of the site generated from CDM databases [110], [111]. This can help to predict that if or not the site is capable of achieving the desired number for patient recruitment.
5.4 Social Media Listening

It is recognized that healthcare organizations, physicians, and patients can get benefited from the use of social media [112]. Data managers being data custodian are comparatively better suited to take advantage of data processing than any other stakeholder in clinical research. Social media websites can be leveraged for site identification, patient recruitment and monitoring of adverse drug events, Figure 5.8.

Potential study sites can be identified based on geography or infrastructure by sentiment analysis of the associated physician; integrating the data with the information mentioned in various trial registries or patient registries [109]. Social media can be helpful to extract health informatics information based on socioeconomic class, age, gender [113] and other major data inputs including signals related to the potential adverse events. CDM can combine and processes the information about existing and emerging health behavior(s), which can be utilized for relabeling or reconfirming the safety profile of the vaccine/drug [114].

Figure 5.8: Social Media Listening, facilitated by CDM databases
5.5 Risk Based Monitoring (RBM)
Global clinical research industry is in the state of transition from SDV (Source Data Verification) to SDR (Source Data Review) for site monitoring. This involves performing most of the task as per the RBM plan from the remote location through centralize monitoring efforts, rather than going to the site to oversee the process. This is not only helpful for identification and mitigation of the risk in a timely manner but is expected to save time and money spent on overall logistics [115].
Role of DMs can be easily evolved and combined with that of clinical trial monitors. DMs can contribute in an effective manner as they have the domain experience of EDC/RDC, which like RBM involves real time data processing. CDM role optimization can provide holistic trial insights resulting in improved data quality and patient safety [116].

5.6 Big Data
Big Data refers to new technologies providing management and processing capabilities, targeting massive and disparate data sets [117]. CDM role may become critical for managing Big-data.
DMs can take-up the challenge by application of efficient practices and tools; as otherwise the processing cost is usually unjustifiable for the organization thus resulting into under-utilization of data [118].
Combining clinical trial data with Big Data, to help forecast trial success or to establish brand positioning [106]; positive insights that are likely to emerge are correct applications of clinical analytics [119].