The CDM Role Evolution

Version 1 (August 2022)

Authors

- Catherine Célingant, Executive Director, Data Monitoring & Management – Oncology, Pfizer
- Patrick Nadolny, Global head, Clinical Data Management, Sanofi
- Inder Sachdeva, Portfolio Delivery Leader, Clinical Data Sciences & Safety, Cognizant
- Ian Shafer, Principal, PwC

Reviewers and Contributors

- Sanjay Bhardwaj, Head of Clinical Technology Strategy & Operations, Abbvie
- Jérémy Nadolny, Sr Business Associate, Trial Operations
- Demetris Zambas, Global Head, Data Monitoring & Management, Pfizer

The Society for Clinical Data Management (SCDM) would like to acknowledge the many volunteers who have contributed to the development of the SCDM reflection papers used as the primary basis for this topic brief.

Methodology

The SCDM Innovation Committee seeks to provide thought-leadership to our industry and support the SCDM vision of "*leading innovative clinical data science to advance global health research and development*". To that end, the SCDM Innovation Committee strives to demystify Clinical Data Science (CDS) and support the development of all Clinical Data Management (CDM) professionals, from subject matter experts (SMEs) working on clinical studies to CDM leaders setting the direction of their organizations.

The SCDM Innovation Committee is publishing topic briefs intended to serve as orientation guides on specific areas which are contributing directly or indirectly to the evolution of CDM into CDS. The content of those topic briefs is primarily an extract from the previously published SCDM Reflection Papers[1,2,3] which collectively provide a cohesive and comprehensive overview of CDS from the point of view of industry leaders. Due to the recent emergence of the CDS discipline and the absence of a comprehensive literature base regarding CDS within the Drug Development industry, this content was gathered from industry leaders through a consensus-based methodology. As CDS matures and technology evolves, we anticipate that literature on this topic will blossom.

Introduction

The human aspect of our changing CDM landscape cannot be underestimated. As new data sources, new study designs and supporting technological advances are taking a foothold in patient centric Drug Development, the role of the Clinical Data Manager is becoming much more complex. This change is at the same time a blessing and a challenge. It offers new opportunities to data experts who have been viewed as the central steward of clinical data quality - the Clinical Data Managers. However, it is a dramatic shift leading to the need to upskill and prepare CDM professionals.

Overall, this is a unique opportunity to re-shape our identity. As such, CDM Leadership and SMEs must proactively guide the transition of their staff from Clinical Data Managers to Clinical Data Scientists. Organizations must consider how to best combine or split responsibilities across role(s) to leverage internal expertise and talents. Additionally, while this already exists in some companies, CDM organizations must consider the best approaches to centralized data monitoring. This will require the use of data review and analytical skills to aid site monitoring teams to efficiently drive onsite and remote site monitoring activities.

Topic brief

In this topic brief, we will evaluate four specific dimensions of the existing and future Clinical Data Manager role: foundational knowledge, expected competencies, soft skills, and best practices. Those four dimensions are setting a generic framework that could be leveraged to define the expectations for any role.

The traditional role of Clinical Data Managers

Even if some companies have started to expand the scope of their CDM responsibilities, **today**, the most common CDM activities are still revolving around the following:

- The lifecycle of data collection tools and clinical data standards
- The end-to-end data flow from collection to archival
- The review and reconciliation of clinical study data
- The management of third-party data and related vendors
- One harmonized set of best practices
- One main CDM tool: Traditional Electronic Data Capture (EDC)
- The project management and documentation for all the responsibilities above

As the industry's leading CDM organization, the SCDM has created anchor points for Clinical Data Managers to successfully meet those expectations through the Good Clinical Data Management Practice[4] (GCDMP©) and the certification program for Clinical Data Managers[5]. Together, they have been defining for almost two decades, the expected competencies, foundational knowledge and best practices of today's CDM world.

First published in 2000, the GCDMP© provides **best practices** supporting CDM organizations in their implementation of high quality CDM processes for paper and EDC-based studies. Its 28 chapters listed in Appendix 1 also guide Clinical Data Managers preparing for CDM training and education.

The SCDM certification program launched in 2004, identifies 70 **expected competencies** organized into 8 domains (see Appendix 2). 85% of the competencies are concentrated in 4 core domains: **Design, Project Management, Data Processing and Programming**.

In addition to the competencies themselves, 25 **foundational knowledge** topics were confirmed by the 2015 and 2018 SCDM Task Analysis Surveys. Those topics include but are not limited to:

- Therapeutic development and clinical research fundamentals
- Scientific method
- Good Clinical Practices (GCP) and other guidance
- Software Development Life Cycle (SDLC) concepts
- Audit methodologies
- Project management fundamentals
- Basic statistical concepts
- Data and metadata models, standards and terminologies including medical terminologies
- Workflow design, analysis, and control fundamentals

Last, beyond the established SCDM competencies, foundational knowledge and best practices, **soft skills** are essential building blocks for Clinical Data Managers. The most commonly expected soft skills are:

- Attention to details
- Logical thinking
- Adaptability
- Ability to articulate complex concepts to the trial teams
- Ability to investigate and troubleshoot complex data trends
- Ability to work with cross-functional teams

While there are some variations across companies, the CDM role framework below, based on core competency domains, foundational knowledge, best practices, and soft skills has long been representing the core expectations from Clinical Data Managers.

Best Practices	**Soft Skills**
Data Management Plan Paper CRF and EDC Lifecycle Vendor Selection and Management External Data Management Assuring and Measuring Quality Patient-Reported Outcomes Medical Coding	Attention to details Logical thinking Adaptability Ability to articulate complex concepts Ability to work with cross-functional teams Ability to troubleshoot complex data trends
Core Competencies Areas	**Foundational Knowledge**
Design Programming Data Processing Project Management	Clinical Development Methodology Regulations Software Development Life Cycle (SDLC) Audit Methodologies Standard Models and Terminologies Workflow Management

Fig 1. The traditional CDM role framework

In summary, CDM is responsible for the lifecycle of clinical data from collection to delivery for statistical analysis in support of regulatory activities. CDM primarily focuses on data collection, data flow and **data integrity** (i.e., ensuring that data is managed the right way).

The role and skillsets of Clinical Data Scientists

CDS expands the scope of CDM by adding the data risk, data meaning and value dimensions for achieving **quality** (i.e., data is credible and reliable). CDS also expands the scope of CDM beyond the study construct by requiring the ability to generate knowledge and insights from clinical data to support other clinical research activities which requires different expertise, approaches, and technologies.

The new world of CDM (i.e., CDS) must be able to continuously improve capabilities across multiple dimensions, including regulatory, operations, technology, and data. Clinical Data Scientists need to rely on effective and efficient processes and controls, enabled by technology to support novel and increasingly complex trial designs, across multiple delivery modalities, all while adapting to evolving global and local regulations. As a result, Clinical Data Scientists need additional competencies to deliver the high quality and high integrity data needed to drive the expected study outcome.

Evolution of the Best Practices

First, Clinical Data Scientists will need complementary best practices framed around the 4 following interconnected major themes covered in the 2022 SCDM Innovation Topic Briefs.

The 5 Vs of clinical Data[6] supporting advanced clinical research and technologies

- Enabling new protocol designs such as adaptive and master protocol
- Enabling the decentralization of clinical trials
- Maximizing Real-World Data (RWD) and Real-World Evidence (RWE)
- Generating secondary data assets (e.g., synthetic control arms)
- Leveraging the reliability and affordability of mobile health including sensors & wearables
- Integrating various data structures through fit for purpose technologies

The evolution of Clinical Data Reviews[7] to ensure data quality

- Adapting to the decline of traditional EDC centricity and increase in data variety
- Focusing on the credibility and reliability of trial results
- Reviewing increasingly complex and study specific data
- Allowing data tagging, exclusion, and curation

The automation of CDM Driven Activities[8]

- Establishing a pragmatic automation roadmap
- Leveraging all components of Artificial Intelligence (AI)
- Maximizing machine to human interactions
- Enabling new capabilities such as identifying data anomalies without explicit validation checks programming

The adoption of Risk based CDM approaches[9] aligned with new regulations including

- Quality by Design (QbD)
- Critical to Quality (CtQ) factors
- Critical data and processes
- Risks lifecycle management including assessment, root cause analysis, etc.
- Ensuring operational feasibility

Evolution of the soft skills

CDM Organization must establish "a culture that **values and rewards critical thinking** and open, proactive dialogue about what is **critical to quality** for a particular study or development programme, going **beyond sole reliance on tools and checklists**" [10]. This fundamental change, which is encouraged by regulators, mandates to move away from our one-size-fits-all approach based on **one** traditional EDC centric data flow and **one** set of sequential processes, predominantly driven by checklists and detailed operational documents.

This represents a mindset change from logical to **critical thinking** which is at the core of our role's evolution. In their lead position, Clinical Data Scientists have to drive the study team through all potential scenarios to optimize operational study execution while minimizing risks to the patients' safety and the reliability of the trials results. This requires broader cross functional collaboration. As an example of a new requirement, Clinical Data Scientists should consider the recruitment strategy and the outcome of the study, country and site feasibility as inputs to the study operational plans. Clinical Data Scientists should tailor data collection systems, data review strategies, and training requirements (at the study, country, site and patient level) to account for the risks associated with the expected patient and site diversity (e.g., geographical, cultural, standard of care, experience, etc.).

Additionally, the nature of issues identified through our risk-based CDM approaches such as advanced predictive analytical capabilities enable root cause analysis and corrective actions of potentially ambiguous study specific scenarios. Clinical Data Scientists must be able to implement and/or recommend alternative and operationally sound solutions (i.e., **being a pragmatic contributor**). These solutions will usually entail adjusting processes to prevent re-occurrence, as opposed to merely correcting individual data points through queries. So, given the dynamic and evolving nature of the clinical research landscape, Clinical Data Scientists have a central role in leading cross-functional teams, as well as **driving and owning complex decision** making.

CDS is not a support function but a key stakeholder, responsible for the most critical asset: the study data. Clinical Data Scientists must therefore demonstrate **influential leadership skills** to shepherd the study team through the new and more advanced CDS process. Doing so requires business acumen, technical capabilities, and the ability to manage continuous change within organizations and teams. In that context, there need to be an increased emphasis on the following soft skills:

- Understanding the points of view of a wider range of stakeholders including sites and patients
- Critically assessing risks and their impacts to determine best mitigation strategies
- Ability to manage ambiguous and unfamiliar study scenarios
- Ability to assess the impact of data changes to the reliability of clinical results
- Articulating complex technological and scientific concepts
- Understanding the rationale and ramifications behind study team decisions

Ultimately, Clinical Data Scientists must be tactful and empathic listeners able to drive consensus around complex scenarios and if necessary, demonstrate decisiveness by taking and owning decisions.

Evolution of the foundational knowledge

Ultimately, to remain effective and relevant, CDM professionals need to become comfortable with emerging technologies, standards, regulations and Clinical Development strategies. As such, ensuring a robust foundational knowledge is essential and has to be one of the priorities of the CDM upskilling journey.

The foundational CDS knowledge could be divided into 5 major categories.

- New protocol designs (e.g., Master, Adaptive)
- Decentralized Clinical Trials (DCT) approaches and technologies
- Risk based methodologies and regulations
- Health care data (RWE/RWD), standard models and terminologies
- Artificial Intelligence (AI) concepts and scope of applicability

Evolution of the competency areas

First and foremost, the core competencies of a Clinical Data Manager must be mastered as CDS expands from them. Some CCDM competencies must be deepened and new one acquired.

Below are important CDS competencies to explore.

- **Risk Management** which must go beyond the typical risks impacting study set-up and database lock timelines. Clinical Data Scientists must identify, prevent, and monitor risks related to protocol design (e.g., size and variety of patient cohorts), study set-up, countries involved, profile of selected sites and any other study specific execution activities.
- **Vendor Oversight** transitioning from managing status and timelines to overseeing quality delivery through data-driven insights on vendor performance, tailored to the services rendered. Clinical Data Scientists will need to strengthen their relationships and understanding of interdependencies with the Clinical Operations, Quality Assurance and Procurement functions to effectively oversee a vendor with a single, integrated voice from the sponsor. Last but not least, Clinical Data Scientist will need to ensure that the vendors are able to implement and adhere to the new CDS practices.
- **Knowledge of Patient Centric Processes and Technologies** including the characteristics of different types of data, such as inpatient and outpatient EHR data, sensors, wearables, imaging, etc. and considering the impact of collecting data through telemedicine and home nursing.
- **Process Management,** including patient and site related processes, to identify and resolve systematic deficiencies or behaviors that lead directly or indirectly to data bias and anomalies.
- **Robotic Process Automation (RPA) and Machine Learning (ML) SDLC methodologies** with ability to help build and test automation to streamline and de-risk routine tasks, as well as ML algorithms to detect patterns of missing data, outliers, trends and/or lack of variability which can be applied as a standard within their therapeutic area or set of studies with the right level of human supervision.
- **Advanced data exploration and interrogation methods** (e.g., non-SQL, R, Python) to interrogate and mine high volumes of data (including audit trails) from a variety of data sources. Clinical Data Scientists must drive the development of tools extracting meaningful insights to detect potentially unreliable data threatening the validity of the trial results.

Impact of the role evolution

The evolution from CDM to CDS summarized in this topic brief results from evolving regulations, technologies, and clinical research approaches. This represents a major shift in focus, not only for CDM but for all clinical research stakeholders. For CDM itself, this leads to the evolution of its competencies, foundational knowledge, best practices, and soft skills requiring the following expectations to be added on top of the existing CDM.

Best Practices	**Soft Skills**
Risk-based CDM approaches *(Incl. Operational feasibility, QbD and CtQs)* Advanced Clinical Data Reviews *(incl. Story telling visualizations, systematic errors detection, Data Tagging, Exclusion and Curation)* Implementation of the 5Vs of Clinical Data *(Incl. Secondary data assets such a synthetic control arms and complex d*ata Integrations*)* Automation of CDM Activities *(Incl. Robotic and Intelligent Process Automation)*	Critical thinking Influential leadership Pragmatism Ability to manage ambiguities Ability to make and own decisions
Core competencies Areas	**Foundational Knowledge**
Risk Management Vendor Oversight Patient Centric Technologies Deployment Process Management RPA and ML Based SDLC Advanced data exploration and interrogations *(e.g., non-SQL, R, Python)*	New protocol designs (e.g., Master, Adaptive) DCT approaches[11] Risk-Based methodologies and regulations Health Care Data (RWD/RWE), Standard Models and Terminologies Emerging Data Structures (e.g., Sensors) Automation and Artificial Intelligence concepts

Fig 2. CDS role framework

While this framework will need to adapt in the coming years with further evolutions in technology and regulations, it could be leveraged as a starting point to support the evolution of the CDM roles into CDS. The **soft skills** and **foundational knowledge** expectations will likely be added to the job descriptions and hiring requirements aligned with each organization strategies.

Furthermore, the need to know how to apply those skills to specific tasks (i.e., **competencies**) aligned with new **best practices** will guide training and up-skilling approaches to enable Clinical Data Scientists to take on new roles. While taking different pathways, many CDM leaders will gradually evolve their organization toward their own tailored CDS future. To initiate such a change management endeavor, they must clearly define their own ultimate destination and value proposition for their organization considering the evolution of the industry toward a digital and patient centric future.

This path will be highly influenced by their **current** company landscape (i.e., starting point) including:

- Size (from small biotech to top 10 pharmaceutical companies)
- Geographical footprint
- CDM roles, scope, and structure (e.g., flat, hierarchical, or matrixed)
- Culture (incl. digital literacy, tolerance to mistakes, agility, silos, innovators vs. followers)
- Merger and Acquisition strategies
- Study team composition
- Cross functional dependencies
- Maturity of Data Driven and Risk-Based approaches
- Technologies (e.g., availability of a metadata repository or not)
- Talent pool
- Emerging interdependent functions (e.g., Start-Up, Design Center, Digital Innovation) and roles (e.g., Head of Clinical Data Science, Chief Digital Officer, Digital Integration Specialist, Trial Innovation Lead)

The evolution from the current CDM state into CDS will vary depending on each organization strategy. The long-term objective of the CDM organization will require a clear change management plan to ensure the current structure does not break while aligning with the overall company objectives.

Additionally, the CDS roadmap and change management plan must integrate aspects such as:

- Human resources strategies: Job classification, career ladders, talent acquisition, compensation, onboarding, training, upskilling, mentoring and evaluation
- CDS operating models (e.g., in-house, outsourcing and FSP models)
- Internal and external stakeholder relationship management
- Organizational change including culture

Conclusion

The evolution toward CDS has started and is unavoidable especially as the COVID-19 pandemic has accelerated the decentralization of clinical trials on a scale never seen before. On the one hand, this has led to stronger support from leadership and regulators. It also removed many of the traditional adoption barriers across all stakeholders. On the other hand, the need to evolve quicker is adding pressure to adapt without much pro-active organizational readiness.

While the speed of change is overwhelming, the opportunity to reshape clinical research is unprecedented. It is therefore crucial to act now and define a strategy enabling our Clinical Data Managers to evolve into Clinical Data Scientists fully equipped to embark on the CDS journey.

CDM must transform into CDS rapidly to emerge as a true clinical research enabler. To seize this meaningful opportunity, CDM leaders must take advantage of the recent changes in the clinical research landscape, the significant investment in DCT related infrastructures as well as the growing maturity of automation technologies. This is a complex task requiring thoughtfulness and a clear strategy mandating evolution of the CDM Roles.

References

[1] SCDM, June 2019, The Evolution of CDM to Clinical Data Science – A Reflection Paper on the impact of the Clinical Research industry trends on Clinical Data Management. Available at https://scdm.org/wp-content/uploads/2019/09/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science.pdf

[2] SCDM, March 2020, The Evolution of CDM to Clinical Data Science (Part 2: The technology enablers) – A Reflection Paper on how technology will enable the evolution of Clinical Data Management to Clinical Data Science. Available at https://scdm.org/wp-content/uploads/2020/03/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science-Part-2.pdf

[3] SCDM, August 2021, The Evolution of CDM to Clinical Data Science (Part 3: Evolution of the CDM role) – A Reflection Paper on the evolution of CDM skillsets and competencies. Available at https://scdm.org/wp-content/uploads/2020/08/SCDM-Reflection-Paper-CDM-Role-Evolution-Part-3-.pdf

[4] SCDM Good Clinical Data Management Practice Guidelines, Available at https://scdm.org/gcdmp/

[5] SCDM Certification program for Clinical Data Managers, Available at https://scdm.org/certification/

[6] SCDM, March 2022, The 5Vs of Clinical Data (Topic Brief), Available at https://scdm.org/clinical-data-science/

[7] SCDM, April 2022, The Evolution of Clinical Data Review (Topic Brief), Available at https://scdm.org/clinical-data-science/

[8] SCDM, June 2022, The Automation of Clinical Data Management Activities (Topic Brief), Available at https://scdm.org/clinical-data-science/

[9] SCDM, July 2022, Risk-Based Clinical data Management (Topic Brief), Available at https://scdm.org/clinical-data-science/

[10] ICH, ICH E8, General Considerations for Clinical Trials Draft Guidelines, Available at https://database.ich.org/sites/default/files/E8-R1_EWG_Draft_Guideline.pdf

[11] SCDM, April 2020, Decentralized Clinical Trials (Topic Brief), Available at https://scdm.org/clinical-data-science/

Main abbreviations

AI	Artificial Intelligence
CtQ	Critical to Quality
CDM	Clinical Data Management
CDS	Clinical Data Science
CRF	Case Report Form
DCI	Data Collection Instrument
DCT	Decentralized Clinical Trials
DMP	Data Management Plan
eCOA	Electronic Clinical Outcome Assessment
EDC	Electronic Data Capture
GCDMP	Good Clinical Data Management Practice
GCP	Good Clinical Practice
ML	Machine Learning
QbD	Quality by Design
RPA	Robotic Process Automation
RWD	Real-World Data
RWE	Real-World Evidence
SCDM	Society for Clinical Data Management
SDLC	Software Development Life Cycle
SME	Subject Matter Expert

Appendix 1 - List of GCDMP Chapters

GCDMP Chapters	
Data Privacy	External Data Transfers
Data Management Plan	Patient-Reported Outcomes
Project Management for the Clinical Data Manager	CDM Presentation at Investigator Meetings
Vendor Selection and Management	Training
Data Management Standards in Clinical Research	Metrics in Clinical Data Management
Design and Development of DCIs	Assuring Data Quality
Edit Check Design Principles	Measuring Data Quality
EDC - Concept and Study Start-up	Data Storage
EDC - Conduct	Data Entry Processes
EDC - Study Closeout	Coding Dictionary Management & Maintenance
CRF Completion Guidelines	Safety Data Management and Reporting
CRF Printing and Vendor Selection	Serious Adverse Event Data Reconciliation
Database Validation, Programming & Standards	Database Closure
Laboratory Data Handling	Clinical Data Archiving

Appendix 2 - List of SCDM Certification domains

Competency domain	# of competencies	% of total competencies
Design: Identification and set-up of all data collection instruments (DCIs) such as EDC and eCOA, data handling and reporting tools leveraging clinical data standards. It also includes core CDM documents such as the Data Management Plan (DMP) and Case Report Form (CRF) Completion Guideline	21	*30.0%*
Project Management: Ensuring oversight of CDM activities from study initiation to study close-out including vendor management	16	*22.9%*
Data Processing: Data Lifecycle from collection to archival. Includes the collection, transfer, import, cleaning, coding, reconciliation and quality assessment of study data	15	*21.4%*
Programming: Creation of the required tools defined during study design. Scope includes programming of the eCRF (Screens and Edit Checks), reports, ad-hoc querying, data imports, transformations and extracts	8	*11.4%*
Testing: Definition and execution of testing strategies for required tools	2	*2.85%*
Training: Ensuring understanding of CDM processes across the organization	2	*2.85%*
Personnel Management: Ensuring CDM staff oversight	3	*4.3%*
Review: Expert review of study and CDM deliverables	3	*4.3%*