

From Data Management to Data Scientist: Navigating the Future of CDM in Clinical Trials

PRESENTED BY:
Debra Anderson
SVP Strategic Partnership



CHANGING ENVIRONMENT OF CLINICAL TRIALS

Decentralization and real time data collection

- ✓ Data streams or feeds as opposed to static data transfers

Risk based monitoring

- ✓ ICH E6 Guidelines
- ✓ Timely, remote data review by trained experts

Perfect data versus Critical data

- ✓ Prioritize critical data that is accurate and fit for purpose
- ✓ Audit trails

Expansion of Clinical Data Sources

- ✓ Wearable Devices and Sensors
- ✓ Real World Data
- ✓ -Omics Data

Changes in Technology

- ✓ Artificial Intelligence / Machine Learning
- ✓ Cloud Based Data Platforms
- ✓ Visualization Tools



KEY DIFFERENCES

Aspect	CDM (Clinical Data Management)	CDS (Clinical Data Science)
Primary Goal	Maintain data accuracy and consistency	Ensure data quality and relevance
Quality Method	Rely on routine quality checks	Embed quality within the design phase
Thought Process	Emphasizes structured, logical outcomes	Prioritizes analytical and critical thinking
Trial Type Focus	Traditional randomized controlled trials	Flexible, adaptive, and master protocols
Key Data Source	Data collected directly from study sites	Data from electronic sources and decentralized trials (DCTs)
Process Approach	Standardized processes across multiple studies	Custom, risk-based processes tailored to each study
Data Volume	Manages relatively low data volumes from fewer sources	Handles large, complex data from multiple sources
Complexity of Data Flows	Simple and straightforward data pipelines	Complicated data interactions and flows
Vendor Role	Oversees basic vendor management tasks	Provides oversight and collaboration with vendors
Data Handling Tasks	Focus on cleaning and correcting data	Review, curate, tag, and manage data exclusions

DATA SCIENCE FUNDAMENTALS



Data Wrangling or Munging

- Cleaning, transforming and organizing data into new structure and usable format
- Handling missing values, correcting data inconsistencies, removing outliers, merging data
- Supports reliable insights when summarizing data



Basics of Machine Learning, AI, and Predictive Modeling

- Be prepared to engage in the discussion and understand what activities are included
- Utilize online options for learning the basics
- Apply concepts to small projects or case studies to gain practical experience



Model Validation

- Evaluating how well a machine learning model performs on unseen data
- Splitting data into training and test sets
- Assessing performance metrics like accuracy and precision

CHECK YOUR MINDSET



Fixed Mindset

Intelligence is static

...avoid challenges



CHALLENGES



...embrace challenges

...give up easily



OBSTACLES



...persist in the face of setbacks

...see effort as fruitless or worse



EFFORT



...see effort as the path to mastery

...ignore useful negative feedback



CRITICISM



...learn from criticism

...feel threatened by the success of others



SUCCESS OF OTHERS



...find lessons & inspiration in the success of others

As a result, they may plateau early and achieve less than full potential.

As a result, they reach ever-higher levels of achievement.

All this confirms a **deterministic view of the world**.

All that gives them a **greater sense of free will**.

SEEK OUT OPPORTUNITIES TO LEARN

Stay Current with Industry Trends:



- Regularly engage with industry publications, webinars, and conferences to keep up with emerging technologies and methodologies
- Subscribe to newsletters or join professional organizations like SCDM to stay informed
- Use this knowledge to identify and implement best practices in your work

Collaborate Across Functions:



- Strengthen communication skills to effectively collaborate with biostatisticians, programmers, and regulatory experts
- Participate in cross-functional meetings to understand the broader goals of clinical trials and how data supports them
- Build relationships that facilitate knowledge sharing and innovation within your organization

SEEK OUT OPPORTUNITIES FOR DEVELOPMENT

Seek Mentorship and Networking Opportunities:

- Connect with professionals who have transitioned into Data Science roles to gain insights into their journey
- Join online forums to network with peers and experts.
- Actively seek feedback and mentorship to identify areas for improvement and guide your career development.

Stay Open to Change:

- Approach new challenges as learning opportunities and remain open to adapting to changes in technology and processes.
- Regularly reflect on your progress and set achievable goals to maintain motivation and focus.
- Embrace continuous learning to keep your skills relevant and impactful in the evolving landscape of clinical trials.

Become a Clinical Data Science Champion within your Organization:

- As you develop these skills share them within your organization.
- Become a change agent.
- Train others and serve as a mentor.

DEVELOP TECHNICAL SKILLS

Gain proficiency in programming languages such Python, R, or SQL

Disclaimer: **Review terms and conditions and privacy agreements before signing up.**

- **DataCamp (<https://www.datacamp.com/>)**
Both R and Python tutorials are available, focusing on data manipulation, visualization, and statistics. While much is paid, several introductory courses are free.
- **Kaggle (<https://www.kaggle.com/>)**
A data science community offering free tutorials on Python and R. Users can access real datasets and compete in challenges.
- **Codecademy (Free Version) (<https://www.codecademy.com/>)**
Offers introductory Python courses to learn the basics

Massive Open Online Courses available

- **edX (<https://www.edx.org/>)**
Offers free courses from universities like Harvard and MIT.
 - Suggested courses:
 - *Data Science: R Basics* (Harvard)
 - *Python for Data Science* (IBM)
- **Coursera (Free Audit Option) (<https://www.coursera.org/>)**
Free access to many courses when you audit.
 - Suggested courses:
 - *R Programming* (Johns Hopkins University)
 - *Python for Everybody* (University of Michigan)

Open Source Projects on GitHub

Open Source Projects on GitHub – Collaborate on Python or R projects to gain real-world experience.

Author Name: Debra Anderson
Company: Veramed

Email:
debra.anderson@veramedinc.com

Website:
<https://www.veramedinc.com/>

