Paper PD01

Bridging Statistical Programming and Data Management: A Job Rotation Experience

Frederikke Høyer, Ferring Pharmaceuticals A/S, Copenhagen, Denmark

ABSTRACT

This presentation explores insights from a statistical programmer's rotation into data management. Through this experience, a holistic understanding of data lifecycle emerged, emphasizing cross-functional collaboration and skill enrichment. The rotation facilitated enhanced collaboration between statistical programming and data management teams, optimizing processes and fostering cultural adaptation. Key takeaways include a broader perspective on data handling, skill enrichment through exposure to diverse tools and practices, and identification of opportunities for process optimization. This journey exemplifies the value of job rotations in fostering cross-functional synergy and driving excellence in data-driven decision-making within pharmaceutical organizations.

INTRODUCTION

Job rotations are a powerful career development tool, allowing us as employees to gain a broader understanding of different organizational functions and enhance our skills. In the pharmaceutical and clinical research sectors, where data plays a crucial role in decision-making, a job rotation between statistical programming and data management offers unique opportunities. This article explores the experience I had as a statistical programmer transitioning into a data management role and the valuable lessons I learned from this cross-functional move.

UNDERSTANDING THE TWO WORLDS: STATISTICAL PROGRAMMING VS. DATA MANAGEMENT

Statistical programming and data management are closely linked, yet they have distinct focuses. Statistical programmers primarily deal with analyzing and presenting data to meet regulatory requirements, ensuring that the data is ready for statistical analysis and reporting. Their work involves transforming raw data into meaningful insights, which inform decision-making in clinical trials and drug development.

Data managers, on the other hand, are responsible for the collection, quality control, and processing of raw data. Their work begins at the start of a clinical trial and continues through to database lock, ensuring that the data collected is complete, accurate, and reliable. While statistical programming focuses on the output of data, data management is concerned with the input and maintenance of data integrity throughout the trial process.

MY JOB ROTATION EXPERIENCE

Taking on a rotation from statistical programming to data management gave me insights into the entire clinical data lifecycle, which enhanced my cross-functional understanding. However, the shift also brought unique challenges, especially with prioritizing tasks between the two roles.

LEARNING THE DATA MANAGEMENT PROCESS

As a statistical programmer, transitioning into a data management role during a job rotation provided me with valuable insights into how the two functions operate within a clinical trial. While

both roles are essential to ensuring high-quality data, they focus on different stages of the trial process, interact with different stakeholders, and follow distinct workflows.

Here's what I learned about the key differences between data management and statistical programming, particularly from the perspective of a programmer.

MAIN DIFFERENCES BETWEEN DATA MANAGEMENT AND STATISTICAL PROGRAMMING FROM A PROGRAMMER'S PERSPECTIVE

FOCUS AND OBJECTIVE

In data management, I focused on the collection and quality of clinical trial data. The primary goal was to ensure the accuracy and integrity of the data from the start of the trial until the database lock (DBL). Working as a DM involves designing databases, setting up electronic data capture (EDC) systems, developing data validation checks, resolving data discrepancies, and performing ongoing quality control of the data. DM focus heavily on ensuring that data is clean and compliant with regulatory requirements long before it reaches the analysis phase.

As a statistical programmer, my objective is very different. I work with cleaned and finalized datasets that I receive after data management has locked the database. My primary task is to transform this data into analyzable formats, applying statistical methodologies to generate tables, listings, and figures (TLFs) for clinical reports and regulatory submissions. While data management ensures data quality, I focus on deriving insights and reporting results that meet statistical and regulatory standards. This process involves writing code (usually in SAS or R) to process data, create analysis datasets following industry standards (e.g., CDISC ADaM), and produce final outputs such as efficacy and safety reports.

CRITICAL STAGES IN THE CLINICAL TRIAL PROCESS

When I stepped into data management, I saw how critical this function is during the data collection phase of a clinical trial. Data managers start being involved in the clinical trial as soon as a protocol outline has been approved, continuously overseeing the data entry process, resolving queries, and ensuring quality. Their work is ongoing and iterative, with frequent interactions with clinical teams to address discrepancies and ensure clean data is available throughout the trial.

In my programming role, I usually step in right before first patient first visit and my work continue after the database lock, when the data is considered final. At this point, the data is ready for final analysis, and my role becomes crucial in transforming it into analysis-ready datasets and producing the outputs required for clinical decision-making and regulatory review. My work typically revolves around fixed deadlines, often tied to submission timelines or interim analyses.

KEY STAKEHOLDERS

In data management, I found myself working closely with clinical teams, EDC vendor and medical monitors to ensure correct design and setup of EDC. In this role, I saw the importance of collaborating with vendors who provide data capture platforms and maintaining a direct line of communication with clinical operations to manage data flow.

In my programming role, my primary collaborators are biostatisticians. I follow their guidance from the statistical analysis plan (SAP) to prepare datasets and create outputs. I also work closely with medical writers and regulatory teams, providing them with the statistical outputs they need to draft clinical study reports or respond to regulatory inquiries. Additionally, once we submit the data, I often have interactions with regulatory agencies, such as the FDA, PMDA or EMA, ensuring that our data packages and analysis meet their strict submission standards.

TOOLS AND SYSTEMS

During my time in data management, I worked extensively with EDC systems, which were new to me as a programmer. These systems are the backbone of data collection in clinical trials.

As a programmer, I primarily use software like SAS analyze datasets. I am responsible for ensuring the data is formatted according to industry standards such as CDISC SDTM and ADaM, which prepare the data for regulatory submission. My focus is not on collecting data but on transforming and analyzing it to produce outputs for decision-making and reporting. I also use version control to manage code and ensure reproducibility of my work.

CHALLENGES FACED DURING THE ROTATION

While the job rotation offered significant benefits, it also came with challenges. I quickly realized that I as statistical programmer had to adapt to the different pace and focus of data management work. Unlike the project-driven nature of statistical programming, data management requires continuous attention to detail and a proactive approach to handling data issues. Additionally, the transition requires a mindset shift, as the focus moves from analyzing outcomes to ensuring the accuracy of inputs.

The downside for me was that different demands of both roles resulted in divided attention. And I was not able to complete neither task with full focus in some cases. For example, while working on an important statistical analysis, data management issues arose, and required my immediate attention, and forced frequent shifts between roles. Data management often requires real-time problem-solving, such as resolving data queries or dealing with discrepancies, which in my case was clashing with programming deadlines. It was challenging to manage the urgent data management requests at the same time as preparing ADaMs for DBL or other high-pressure timelines.

Managing two sets of responsibilities, each with its own deadlines and expectations, can lead to burnout. The constant need to prioritize and switch between tasks can create the feeling of stress and fatigue, making it difficult to maintain high performance in both roles.

PLANNING FOR A JOB ROTATION

There are no specific requirement or agreement within my organization on how to handle a job rotation and it was therefore up to me to define what I wanted to learn and the outcome. When you only have a limited number of months to learn a specific role it is crucial to set some measurable and accountable short term goal and have it aligned with your nearest manager.

During the planning phase I also noted that a job rotation might not be the line managers dream or need. Actually, it can be a challenge for them to find relevant tasks for a person who is only available for a limited time since many trial related activities are spanning over a long period of time. It is difficult to have a successful job rotation if you do not have the support from the management of the new role.

CONCLUSION

As a statistical programmer, transitioning into data management helped me understand the **end-to-end process** of clinical trial data. Data management focuses on ensuring the **quality and integrity** of the data during the collection phase, while statistical programming is centered on analyzing the cleaned and final data and producing **regulatory-compliant outputs**. The two roles

have distinct stakeholders, tools, and processes, but they are both critical to the success of a clinical trial. However, the challenges of managing task prioritization between the two roles required careful time management, communication, and strategic planning. Despite the difficulties, understanding these differences has allowed me to appreciate the importance of collaboration between data managers and programmers and improved my ability to anticipate and address data issues early in the process.

CONTACT INFORMATION

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

Author Name: Frederikke Høyer
Company: Ferring Pharmaceutical A/S
Email: Frederikke.hoeyer@ferring.com