PharmaSUG 2025 - Paper LS-093

Bridging the Gap: Leadership of Statistical Programmers in Clinical Trials

Allison Covucci, Bristol Myers Squibb
Xiaohan Zou, Bristol Myers Squibb

ABSTRACT

In the complex environment of clinical trials, various functions collaborate to deliver more medications to more patients faster, effectively, and efficiently—from protocol development, clinical trial startup, data collection, and database lock, to submission to Health Authorities (HA), disseminating scientific information, and ensuring the accessibility of new therapies to the public. Each function often operates from its own perspective, without a comprehensive understanding of how and where the data they handle will be utilized to support various objectives through statistical analyses. This is where the role of statistical programmers becomes pivotal. As the final users of the data, they possess a unique vantage point that allows them to see the entire process from data collection to statistical analyses. This presentation highlights the critical leadership qualities that statistical programmers possess to effectively act as a bridge between various functions. Statistical programmers work closely with statisticians to ensure that the data collected is effectively transformed into meaningful statistical results. They engage proactively during the development of the protocol, eCRFs, and CRF completion guidance to ensure the collection of relevant and necessary data for effective clinical review and statistical analysis. They collaborate with data management to address data issues, with clinical teams to combine data analytics with the clinicians' medical insights, and with regulatory teams to ensure regulatory compliance. Throughout the entire process, statistical programmers streamline processes, develop standards, introduce automation, push boundaries, and drive successful clinical trials.

INTRODUCTION

Clinical trials are essential for the development of new medications, ensuring that they are safe and effective for patients. They encompass multiple stages, including protocol development, clinical trial startup, data collection, database lock, submission to Health Authorities, dissemination of scientific information, and ensuring the accessibility of new therapies to the public. However, the complexity and multifaceted nature of clinical trials present challenges. Functions within this process often operate from their own perspective, focusing on specific tasks without a comprehensive understanding of how their data will be utilized to support various objectives through statistical analyses.

Statistical programmers, as the final users of the data, have a unique vantage point that allows them to observe the entire process from data collection to statistical analyses. They play a crucial role in bridging the gap between various functions, ensuring that data is effectively transformed into meaningful statistical results. In addition, their involvement is pivotal in streamlining processes, developing standards, and introducing automation.

This paper explores how statistical programmers act as a bridge between various functions. It will also delve into the key leadership qualities of statistical programmers that drive the success of clinical trials, as well as the challenges they face in this complex and dynamic environment.

THE ROLE OF STATISTICAL PROGRAMMERS

Statistical programmers play a crucial role in clinical trials by supporting the development, regulatory approval, and market acceptance of new medications through their comprehensive analytical and clinical data expertise. They are responsible for designing, developing, and implementing statistical programming solutions for integrating, analyzing, and reporting clinical data to support various objectives through statistical analyses.

Statistical programmers work closely with statisticians to ensure that the data collected is effectively transformed into meaningful statistical results. They collaborate with cross-functional development teams in various capacities: contributing to the development of protocols, electronic Case Report Forms (eCRFs), and CRF completion guidance; addressing data issues with data management; combining data analytics with medical insights alongside clinical teams; and ensuring regulatory compliance with regulatory teams. By streamlining processes, developing standards, introducing automation, pushing boundaries, and driving the success of clinical trials, statistical programmers play a pivotal role in the overall process.

ENSURING DATA COLLECTION, QUALITY, AND INTEGRITY

Designing a CRF is a critical step in clinical trials, as it ensures the accurate and efficient collection of data. Key CRF design principles include clarity, simplicity, logical flow, consistency, relevance, and minimizing effort at the clinical sites. Additionally, CRF designers focus on ensuring compliance with CDISC standards. However, CRF designers may not always fully understand or be aware of the impact of specific data collection requirements on downstream statistical analysis. This is where statistical programmers provide significant value; they know the impact of data collection on downstream statistical analysis and can provide that perspective to the CRF designers. For example, when implementing programmatic Data Cut Off (DCO) for interim analyses in ongoing trials linkages among CRFs and data domains are critical. Additional fields may be required to connect CRFs and domains to ensure consistency in data points in the post-cutoff database after the programmatic DCO. In this situation, it is important for statistical programmers to proactively contribute to the CRF design process and communicate these requirements to the CRF designers. By doing so, they can ensure that the CRF captures all relevant data needed for accurate and efficient statistical analysis.

CRF completion guidance (CCG) is an essential document in clinical trials. It is critical for ensuring that various personnel involved in data collection and analysis have the same understanding of the exact meaning of data points entered in the CRF. This shared understanding ensures that clinical sites can reliably provide data, data managers and clinical teams can effectively review the data, and analysis datasets can be derived accurately for performing statistical analysis. Statistical programmers play a significant role in the development of CCG by providing insights into how and where the data will be used in statistical analysis and translating this information into the collection requirements of data points in the CCG. They identify and address ambiguities and gaps in the CCG, ensuring that accurate, complete, and consistent data is collected from clinical sites through the CRF. By clarifying data field definitions and providing detailed instructions, statistical programmers help connect the dots and fill the gaps, facilitating reliable data collection and appropriate use of the data in analysis. For instance, when deriving the last known alive date, it is important to make clear in the CCG whether the date recorded in the CRF represents the date the patient was confirmed alive or the date the site contacted the patient's family or friends, which may not necessarily be the patient themselves. This level of detail ensures that the data collected is accurate and meaningful for statistical analysis.

Managing the quality of critical endpoints in clinical trial data is essential for ensuring the success of the studies, which includes achieving accurate and reliable results, maintaining regulatory compliance, ensuring patient safety, producing scientifically valid outcomes. Statistical programmers understand which data points are critical for statistical analysis and can assist in determining the priority of data cleaning efforts. This knowledge is essential for data management (DM) and clinical teams, as it helps them focus on resolving the most critical data issues. Additionally, statistical programmers can help design and implement edit checks and other specific data review processes to identify and resolve data issues early in the trial. For example, by incorporating potential data issues that impact regulatory compliance into routine data review processes, statistical programmers help ensure that these issues are addressed well before the database lock. This proactive approach minimizes the rush to resolve data issues at the last minute.

LEADING EFFECTIVE STATISTICAL ANALYSIS THROUGH COLLABORATION AND EXPERTISE

Statisticians and clinical teams are two key stakeholders for statistical programmers. Their collaboration plays a crucial role in achieving effective statistical analysis by transforming collected raw data into clinically meaningful results. The clinical teams bring the scientific knowledge related to the medical and clinical aspects of the study, while statisticians contribute their expertise in statistical theory and methods.

Together, with the technical skills of statistical programmers, they integrate the scientific knowledge with analytical techniques to produce a thorough and precise representation of the data, which is both scientifically rigorous and clinically relevant.

Statisticians and statistical programmers each bring unique expertise, forming a strong analytical team when combined. Statisticians possess deep knowledge of statistical theory and methods, applying rigorous statistical techniques to analyze data scientifically. In contrast, statistical programmers have thorough knowledge of data collection processes, various databases, and technical skills, including proficiency in statistical programming languages, developing thoughtful, clear, and robust algorithms, and designing analysis datasets that are well-organized, optimized, and structured in a way that facilitates easy and effective statistical analysis. One point of utmost importance is the engagement of both roles in understanding the objectives of the analysis as defined in the protocol and Statistical Analysis Plan (SAP) through discussion and alignment. This engagement leads to the statistical programmer performing the statistical analysis in partnership with statisticians, resulting in a thorough and precise representation of the data. An example of this collaboration is the implementation of the estimands framework. Statisticians define the estimands framework through the protocol and SAP. Statistical programmers then engage and align with statisticians to implement the framework at the data collection level and ensure it is integrated with data standards. They make decisions on how to collect intercurrent events, manage data flow, and create specific analysis datasets that ensure traceability of intercurrent events and corresponding strategies, leading to efficient analysis. Both roles engage in continuous communication to address any issues, refine the analysis, and ensure accuracy. Ultimately, the programming skills, tools, and solutions provided by statistical programmers are utilized to innovatively deliver the final results.

We have been seeing notable successes when there is strong engagement between statistical programmers and the clinical team. This collaboration ensures that the statistical programmers' deep understanding of the data, attention to detail, logical mindset, and analytical skills are effectively combined with the clinical team's medical insights. This synergy ties the science to the analysis, enabling the team to perform clinically meaningful analyses and interpret the results with a deeper understanding. Statistical programmers bring invaluable expertise to the table. They possess thorough knowledge of data collected in clinical trials, understand the relationships among data points in the database, and have an analytic mindset to translate the clinical questions posed by the trial objectives into programmable algorithms. Because they are intimately familiar with the data, statistical programmers can determine whether the instructions can be converted into programmable algorithms, identify any missing pieces, and pinpoint where precise instructions are lacking. For example, consider a tumor criterion that requires information on whether there has been an increase in corticosteroid dosing since the last brain scan. Since corticosteroids are administered daily, the statistical programmer understands the intricacies of data collection in the database and can raise important questions about how to determine an increase in dosing since the last scan. For instance, should the increase be defined as occurring once, several times, or on all days since the last scan? These questions require the clinical team to provide specific clinical insights. This ability to identify and address such detailed questions is one of the strengths that statistical programmers bring to the table. Their expertise helps bridge the gap between raw data and actionable clinical insights, ultimately contributing to the success of the clinical trials.

DRIVING SUBMISSION SUCCESS THROUGH CROSS-FUNCTIONAL COLLABORATION

While the number and complexity of deliverables for health authority submissions have been increasing, along with the growing need to comply with stringent submission guidelines and standards in recent years, statistical programmers play a crucial role in ensuring that the submissions meet regulatory requirements. They possess substantial knowledge of global health authority guidelines, including those of the FDA, PMDA, NMPA, EMA, and others. Collaborating closely, statistical programmers and regulatory teams provide the framework for submission-ready electronic components and outlines of dataset-related standards necessary for successful submissions to health authorities worldwide. In addition, statistical programmers develop reviewer's guides that facilitate health authorities in navigating e-submission packages. These guides offer clear and comprehensive documentation that explains the structure, content, and organization of the datasets, making it easier for reviewers to quickly locate and understand the information they need. When new guidelines are released, statistical programmers are pivotal in swiftly implementing them to ensure compliance. They collaborate with regulatory teams and other functions to

interpret the new requirements, update existing processes, and integrate the standards into e-submissions. By staying current with evolving guidelines and adapting quickly to changes, statistical programmers help ensure that submissions align with the latest regulatory expectations, increasing the likelihood of a smoother review process and timely approvals.

An exemplary demonstration of this leadership is seen in the Real-Time Oncology Review (RTOR) process, an initiative by the FDA to expedite the review of oncology drugs. RTOR enables the FDA to review data earlier than in the traditional submission process, requiring careful preparation and close cross-functional collaboration. In this context, statistical programmers are instrumental in ensuring that submission packages are prepared and submitted early, facilitating a more efficient review. They work proactively with statisticians, medical writers, and regulatory teams to define timelines, identify potential issues that could disrupt the process, and minimize the need for modifications after the initial RTOR submission. Additionally, statistical programmers respond to FDA requests promptly and thoroughly after RTOR submission, ensuring that the review process remains on track and avoiding delays. Through their leadership and expertise, statistical programmers help ensure compliance with regulatory requirements, contributing to the success of the RTOR process.

Through their expertise and proactive collaboration, statistical programmers effectively bridge the gap between functions, driving submission success.

STREAMLINING PROCESSES AND DEVELOPING STANDARDS

Streamlining processes and developing standards are fundamental to the success of clinical trials. These practices not only enhance the efficiency and speed of trial execution but also ensure consistent quality, regulatory compliance, and, ultimately, the successful delivery of new therapies to patients. Statistical programmers are uniquely positioned to drive improvements in both process and efficiency due to their specialized skill set and holistic understanding of the trial lifecycle. Their deep knowledge of complex data and regulatory requirements, combined with a comprehensive view of the entire process—from data collection to statistical analyses—allows them to identify opportunities for optimization.

Statistical programmers work closely with stakeholders across functions—statisticians, clinical teams, data management, and regulatory affairs—to create analysis standards that boost efficiency and increase productivity. These standards provide a foundation for advanced analysis techniques, enhance regulatory confidence, and ensure compliance with evolving guidelines. They allow teams to focus more on understanding the nuances of the study and interpreting its results. Additionally, these standards influence upstream processes, such as data collection, and downstream activities, such as data review, creating a more seamless flow throughout the trial.

By leveraging their deep knowledge in clinical trials, regulatory compliance, and programming solutions, statistical programmers streamline processes, making it quicker and easier to produce accurate results. They develop and implement automation solutions that reduce the need for manual intervention, thereby boosting productivity and minimizing the risk of human error. For example, using Python to generate subject-level data line listings in PDF format that comply with FDA requirements for BIMO has significantly reduced manual work; developing a comprehensive macro library to generate ADaM datasets and outputs to support clinical study reports, submissions, and ad hoc analyses. Through these innovative methods, statistical programmers not only address the immediate needs of clinical trials but also contribute to the advancement of best practices and technological progress in the field.

DISCUSSION

In recent years, the design of clinical trials has become increasingly complex. This complexity is driven by the need for more precise and personalized treatments, the integration of advanced technologies, and the necessity to address diverse patient populations and endpoints. At the same time, the range of therapeutic options has significantly expanded with the development of novel treatments such as targeted therapies, immunotherapies, gene therapies, and personalized medicine. Additionally, clinical trials now involve large and complex databases from multiple sources, including genomic, imaging, and real-world data. Furthermore, the number and complexity of deliverables for health authority submissions have been increasing, along with the need to comply with stringent submission guidelines and standards.

These advancements offer more tailored and effective treatment options, ultimately improving patient outcomes and quality of life. However, they also present significant challenges and complicate the design, data collection and cleaning, statistical analysis, and outcome interpretation in clinical trials. In this situation, more effective and closer collaboration among different functions becomes increasingly critical, and bridging the gaps is essential to ensure successful collaboration. The increasing complexity also introduces pressure on timelines, resources, and regulatory compliance, creating an environment where innovation and adaptability are more essential than ever.

The unique position of statistical programmers in clinical trials allows them to actively engage in the entire process—from data collection to statistical analyses—and to act as connectors among functions. However, statistical programmers must contend with stereotypes within the pharmaceutical industry. They are sometimes perceived as primarily focused on technical tasks and data manipulation, rather than being seen as integral contributors to strategic decision-making or clinical insights. Additionally, they are sometimes viewed as followers who stand behind statisticians, executing tasks specified in the SAP and ADaM specifications, rather than as proactive problem-solvers and innovators.

So, where do statistical programmers stand today in terms of connecting the dots? Are they already bridging the gaps and not being recognized for it, or are they connecting some dots but not speaking up enough to ensure their contributions are acknowledged? Or, perhaps, they are not yet fully stepping into the role of connecting the dots across functions as they should? It's likely a mix of these possibilities. In any case, statistical programmers have the potential to play a much more visible and integral role in clinical trials through their impact on data strategy, automation and innovation. By being open to sharing their insights and perspectives, even when it involves taking risks, statistical programmers can shift from being perceived as followers to proactive leaders in clinical trial strategy. Taking ownership of their role allows statistical programmers to drive the agenda forward with a clear vision and purpose. This ownership also includes being involved from the very beginning of the study, during protocol development, and contributing to the design of eCRFs and data collection strategies, where their expertise can have a long-term impact. Being proactive can help in anticipating challenges and opportunities, ensuring that they are always a step ahead in the process. This approach involves not only addressing current issues but also foreseeing future needs and preparing accordingly. By developing a strong understanding of both the clinical and regulatory landscapes, statistical programmers can add significant value in these early stages, making them indispensable contributors to the overall success of clinical trials.

Moreover, statistical programmers can strive to be integral members of the solution by actively contributing to cross-functional teams. Effective communication skills are key to collaborating with statisticians, data managers, clinical, and regulatory teams. By fostering strong relationships and building trust across these functions, statistical programmers can help ensure that data is accurately transformed into meaningful statistical results that support the overall objectives of the clinical trial. The ability to explain complex technical details in a clear and strategic way is a leadership trait that can significantly elevate their role.

With the advent of AI, the role of statistical programmers is evolving. If programmers are only known as coders, they risk being replaced by AI. However, by positioning themselves as connectors who actively engage in the process and contribute to strategic decision-making, statistical programmers can leverage AI to enhance their skillset. AI can assist in automating routine tasks, allowing programmers to focus on more complex problem-solving and innovative solutions. By embracing AI and integrating it into their workflows, statistical programmers can further solidify their role as indispensable members of the clinical trial team, driving advancements and ensuring the success of clinical trials. Moreover, as AI continues to evolve, statistical programmers who embrace its potential can contribute to the development of AI-driven methodologies that will positioning themselves as leaders in shaping the future of clinical trials.

CONCLUSION

Functional teams may have their boundaries, but data knows no such limits. As the ultimate users of the data, statistical programmers are uniquely positioned to transcend these boundaries, seamlessly navigating through the complexities of data across different functions. Rather than restricting themselves to a narrow, technical role, statistical programmers can thrive by adopting a broader, more holistic view of the clinical trial process. This broader perspective allows them to identify opportunities for improvement and innovation at each stage of the trial.

Statistical programmers possess a rare combination of technical expertise, proactive problem-solving abilities, and a deep understanding of the entire clinical trial process. Their ability is vital for ensuring that clinical trials are executed efficiently and accurately. By strategically leveraging these strengths, statistical programmers don't just support the trial—they drive it forward, ensuring better outcomes and more streamlined processes.

In doing so, statistical programmers significantly enhance the overall success of clinical trials, paving the way for the development of more effective, personalized treatments. Through their leadership, collaboration, and innovative thinking, statistical programmers can play a pivotal role in advancing the future of clinical trials and, ultimately, improving patient outcomes worldwide.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to everyone who contributed to this research, directly or indirectly. Your support and assistance have been invaluable.

CONTACT INFORMATION

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

Allison Covucci
Bristol Myers Squibb
Allison.Covucci@bms.com

Xiaohan Zou
Bristol Myers Squibb
Xiaohan.Zou@bms.com