PharmaSUG 2025 - Paper SI-359

An Introduction to the Role of Statistical Programming in Medical Affairs

Nagadip S Rao, Alnylam Pharmaceuticals, Inc.

ABSTRACT

Medical Affairs (MA) is a specialized division within research-based pharmaceutical companies, dedicated to scientific communication, education, and ensuring the safe and effective use of products. MA plays a pivotal role in supporting the commercialization process and overseeing post-market activities. The MA team comprises interdisciplinary professionals, with statisticians and statistical programming teams being integral to its operations. This paper provides an overview of MA's functionality within the industry, highlighting the role of statistical programming. Additionally, it discusses the nature of deliverables, challenges faced, and presents a real-life example of how statistical programmers support MA in achieving its objectives.

INTRODUCTION

Medical Affairs (MA) is a critical component within research-based biopharmaceutical companies, operating in tandem with the Research & Development (R&D) and Commercial departments as a key strategic function. While R&D is dedicated to the development of new drugs and the Commercial team focuses on marketing and sales, MA is tasked with generating and disseminating data. This data is essential for healthcare professionals (HCPs), payors, policymakers, and other stakeholders to make informed decisions that ensure the optimal use of treatments for patient benefit. Consequently, MA is indispensable in providing the scientific evidence and insights necessary to guide clinical practice effectively (Liliana Virzi, 2023).

Figure 1: Medical Affairs and its primary activities in research-based pharmaceutical company

In a typical research based pharmaceutical company MA group (Figure 1) is primarily responsible for the following 4 activities (Anupma Dhanda Farrington, 2023).

1. Providing impartial, evidence-based scientific and medical information to healthcare professionals, scientific leaders, patient advocacy groups, payors, policymakers, and other stakeholders within the healthcare ecosystem.

2. Gathering insights from various external sources, such as healthcare professionals, opinion leaders, advisory boards, and patient advocacy groups, to inform the organization's decision-making in areas like education, research, development, compassionate use, publications, and strategy.

3. Creating new data on marketed and emerging treatments using Real-World Evidence (RWE), Health Economics and Outcomes Research (HEOR), Investigator Sponsored Studies, or pre- and post-approval studies. This data can support product registration or be non-regulatory in nature

4. Collaborating with industry leaders from various functions, such as R&D, Commercial, and Business Development, to shape the organization's strategic direction for the benefit of patients.

We will explore the role of statistical programming in supporting the above-mentioned MA activities.

ROLE OF STATISTICAL PROGRAMMING IN MEDICAL AFFAIRS

We statistical programmers are skilled in using statistical software to analyze clinical trial data. Their expertise is crucial for transforming raw data into valuable insights that inform medical decisions, shape treatment protocols, and advance scientific knowledge. Within Medical Affairs (MA), statistical programming often involves working with data from various sources that may not always adhere to industry standards (Figure 2). We support medical affairs by creating credible, high-quality outputs for interpretive or exploratory publications, post-approval safety studies, reimbursement support, and marketing/commercial activities (Jangili, 2024).

Figure 2: Medical Affairs Statistical Programming Group (MA-SPG) support activities

The Medical Affairs Statistical Programming Group (MA-SPG) plays a crucial role in supporting various Medical Affairs (MA) objectives. This section discusses the scope and role of MA-SPG, particularly in Real World Evidence (RWE) studies, analysis, and reporting.

Real world evidence study, analysis, and reporting

MA-SPG is involved in supporting prospective observational studies of patients with specific medical conditions who are on therapies of interest. These studies, also known as registry studies, are primarily conducted to gather data and gain insights into long-term trends within specific populations. Registry studies serve multiple purposes, including publications and as source data for Post-Authorization Safety Studies (PASS) and Pregnancy Safety Studies (PSP).

Registry studies typically rely on real-world data gathered from routine clinical practice or observational settings, rather than data collected specifically for research. This approach helps researchers understand how treatments or interventions perform in real-world conditions, outside the controlled environment of clinical trials. By drawing data from diverse sources, registry studies can include large sample sizes, offering a more comprehensive understanding of the outcomes and effects of specific treatments or interventions. A key advantage of registry studies is their ability to encompass diverse populations, mirroring the broader patient demographics seen in clinical practice, thus enhancing the generalizability of the study findings. Often exploratory in nature, registry studies aim to generate hypotheses or insights rather than test specific hypotheses. They can uncover patterns, associations, and trends that may require further investigation in controlled clinical trials.

In May 2019, the FDA issued draft guidance for its Real-World Evidence (RWE) Program, which aims to assess the use of real-world data in regulatory decisions. This has led pharmaceutical companies and medical communities to increasingly recognize the importance of patient registries—large, noninterventional studies—in providing evidence of treatment effectiveness and safety in clinical practice. Registry studies are crucial for evaluating the effectiveness and safety of treatments or interventions in real-world settings, complementing the efficacy data from controlled clinical trials. They are often used for post-market surveillance of drugs, devices, or interventions to monitor their safety and effectiveness once they are approved and widely used (E Dawn Flick, 2023). Various types of registries exist, including general surveillance registries, pregnancy registries for specific drug exposures, and studies on the exposure to medications and devices like pacemakers. A detailed discussion on the nature and outcomes of registry studies is beyond the scope of this paper.

Role of statistical programming in registry studies

The Medical Affairs Statistical Programming Group (MA-SPG) plays a crucial role in extracting and analyzing data from registry studies. MA-SPG utilizes this data for various Medical Affairs deliverables, including post-approval safety surveillance commitment reports, risk management plans (RMP), drug safety surveillance submissions such as post-approval pregnancy safety surveillance reports (PSP), reimbursement submission support, and Periodic Benefit-Risk Evaluation Report (PBER) updates. Additionally, MA-SPG produces reports on the real-world effectiveness of approved drugs of interest and conducts exploratory analyses to gain insights into the safety and efficacy of drugs in real-world settings.

Data from registry studies can originate from various sources with considerable variation in collection standards, often resulting in incomplete or invalid data. MA-SPG faces the challenging task of collaborating with statisticians, data management, and clinical teams to make the best use of available data. Many MA analyses related to registry studies use raw data directly without the intermediate step of mapping into Study Data Tabulation Model (SDTM) standards. This approach limits the utilization of standard macro ecosystems, which are generally focused on generating analysis datasets or Tables, Listings, and Figures (TLF) for Clinical Study Report (CSR) deliverables.

Publication and medical communication

Publications and medical communication are essential for educating healthcare providers (HCPs) and other stakeholders in healthcare management, enabling them to deliver the best possible care to patient populations who can benefit the most. The Medical Affairs Statistical Programming Group (MA-SPG) plays a vital role in preparing manuscripts for prestigious journals such as The Lancet and the New

England Journal of Medicine (NEJM). As statistical programmers, our primary focus is to ensure that the data and analyses presented are not only accurate and highly relevant but also accompanied by compelling visuals.

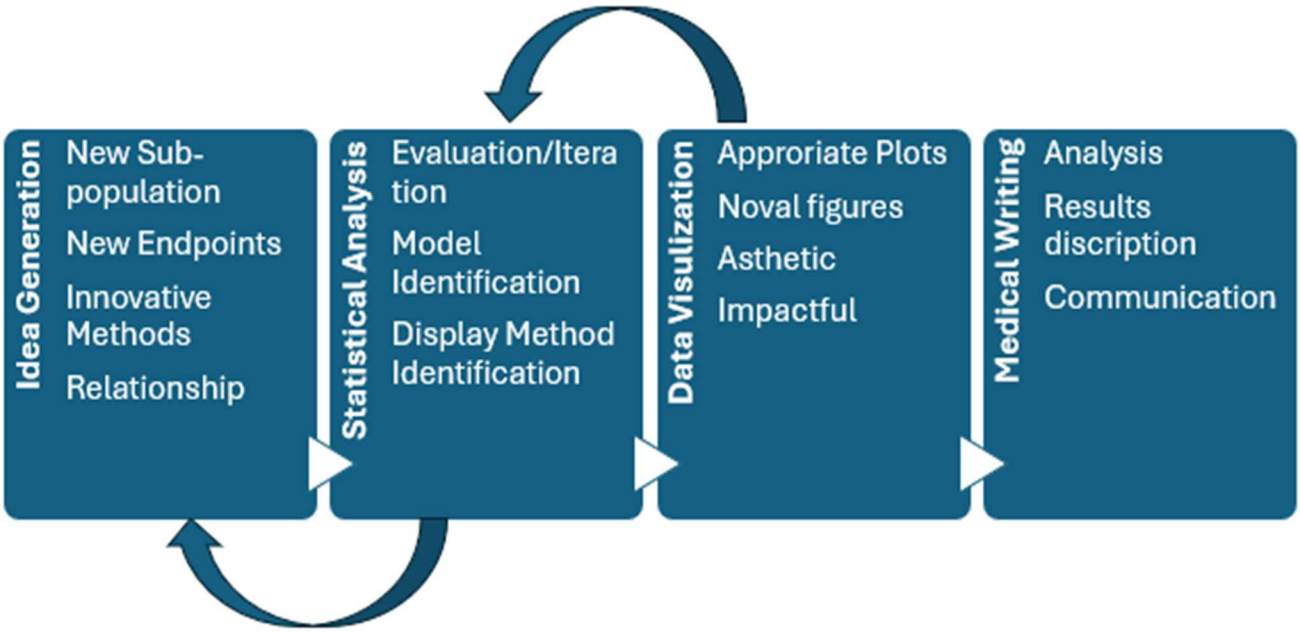

Figure 3: Manuscript and publication processes in medical affairs

A typical publication begins with an interdisciplinary team of clinicians, healthcare research professionals, biostatisticians, and other subject matter experts within the Medical Affairs (MA) team defining an exploratory question of interest (Figure 3). For example, they might investigate whether a subgroup of subjects has a statistically reduced risk of cardiac events over three years compared to the overall study population on the therapy of interest. Statisticians collaborate with programmers to conduct a preliminary feasibility analysis, exploring the therapy's benefit across different patient subgroups. During this phase, various statistical analysis methods are employed to identify the most appropriate approach.

Once preliminary results are discussed and evaluated by the MA team, a decision is made on whether to proceed with the publication. Upon deciding to move forward and identifying target journals, an MA Statistical Analysis Plan (MA-SAP) is developed by the MA statistician and reviewed with the MA Statistical Programming Group (MA-SPG). Unlike a Clinical Study Report (CSR) SAP, the MA-SAP is a more dynamic document that evolves based on insights from analysis results. This flexibility allows for the refinement of the analysis, leading to more robust and reliable findings.

Data Visualization

Data visualization is crucial in clinical publications, transforming complex data sets into easily interpretable visual formats (Nehad A Abudiyab, 2022). This process aids researchers, clinicians, and policymakers in quickly understanding key findings and trends, thereby facilitating better decision-making. Common visualization techniques include bar charts, line graphs, scatter plots, and heat maps, each serving a specific purpose based on the data's nature. For example, survival curves are often used in clinical trials to depict patient survival over time, while forest plots are employed in meta-analyses to show the effect sizes of different studies. Additionally, 'non-traditional' graphics such as Sankey plots like figure 4 (Heidi Taipale, 2025) are used in Medical Affairs (MA) publications to convey messages more clearly and impactfully.

Effective data visualization not only enhances the clarity and impact of the research but also ensures that the information is accessible to a broader audience, including those without a strong statistical background.

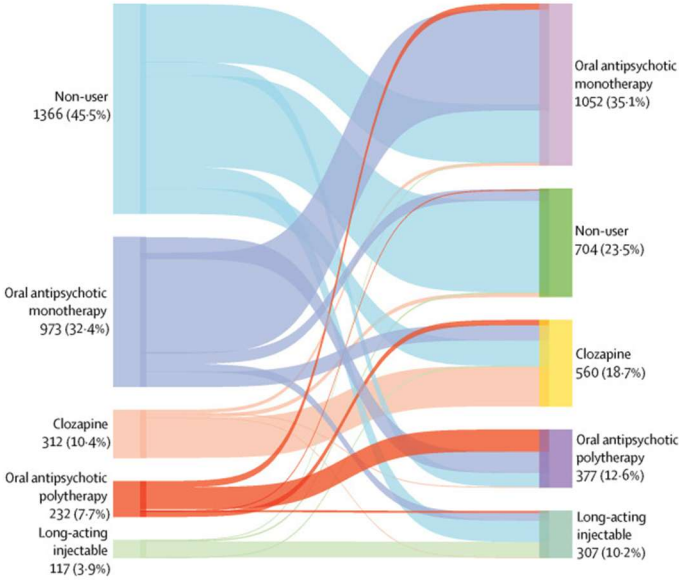

Figure 4: Sankey Plot showing treatment strategies before and after first relapse in 3000 individuals with first-episode schizophrenia (Heidi Taipale, 2025)

In clinical publications, the accuracy and integrity of data visualizations are paramount. Misleading or poorly designed visuals can lead to incorrect interpretations and potentially harmful clinical decisions. Therefore, adhering to best practices in data visualization is essential, such as maintaining appropriate scales, using clear labels, and avoiding unnecessary embellishments that can distract from the data's true meaning. Transparency in the methods used to generate visualizations, including any data preprocessing or statistical adjustments, is also critical for reproducibility and trustworthiness. As the volume of clinical data continues to grow, the ability to effectively visualize and communicate this information will remain a vital skill for MA groups and MA Statistical Programming Groups (MA-SPG).

One of the key aspects of working in the MA group from a statistical programming perspective is data visualization. The adage "a picture is worth a thousand words" aptly describes how complex results and their interdependencies can be conveyed graphically in a much more effective and engaging way than through written descriptions. It is important to remember that our tables and figures are generally catered to a broader audience rather than just for Clinical Study Reports (CSR), and hence they need to convey the message clearly.

A MEDICAL AFFAIRS STUDY AND RELATED MA-SPG DELIVERABLES

To better illustrate role of MA-SPG in the realm of medical affair, an overview of a long-term post-authorization observational study conducted by Alnylam pharmaceuticals to evaluate the long-term safety and effectiveness of real-world use of approved therapy to treat Hereditary transthyretin-mediated amyloidosis (ATTRv). ATTRv is a rare, serious and life-threatening disease that affects multiple systems in the body and is caused by certain genetic mutations.

This registry study is a prospective, multi-national, observational study designed to document the natural history, clinical characteristics, and management of ATTR amyloidosis in routine clinical care. This study involves utilizing secondary data over a 10-year period to assess the safety of a novel RNAi based therapy

in ATTRv amyloidosis patients under real-world conditions (clintrial.gov, n.d.). Specifically, it will collect selected events of interest in a cohort of patients diagnosed with ATTRv amyloidosis to meet post-authorization safety study (PASS) and pregnancy surveillance program (PSP) requirements for these treatments along with multiple scientific publications/presentations planned utilizing collected data.

Data Collection

All data collected was from medical records collected during patient visits and was also extracted from available EMR's. No extra visits, tests, or procedures were needed for this study. Collected patient data was entered into the EDC system, which was cleaned and provided to MA-SPG periodically. Some of the key data elements collected and entered into study database include demographics data, ATTR diagnosis and related family history and treatment, effectiveness assessment information collected during periodic visits included patient-reported outcomes (PROs) like Kansas City Cardiomyopathy Questionnaire (KCCQ), health care provider assessments of polyneuropathy and cardiomyopathy (Polyneuropathy Disability, NYHA classification) and safety assessment data related to AE's of special interest, SAE's, hospitalization etc. Considering the nature and source of data collected for the registry study as a first step, MA-SPG creates a data completeness report. This report provides an overview of collected data and helps medical affairs groups plan scientific publications/presentations along with other clinical deliverables.

Data Completeness Reports

A data completeness report in clinical registry studies is a document that assesses the extent to which all required data points have been collected and recorded accurately. A sample shell of the data completeness report is shown in Table 1. This report is crucial for ensuring the quality and reliability of the data used in clinical research. Here are some key aspects:

Data Accuracy: Ensures that the data entered the registry is correct and matches the source documents.

Data Timeliness: Evaluates whether the data is collected and entered within the required time frames.

Data Consistency: Checks for uniformity in data entry across different sites and over time.

Data Completeness: Measures the proportion of missing data points and identifies any gaps in the data collection process

	Pre-Enrollment N=x	At Enrollment Date N=x	0-6 months post Enrollent date N=x	6-12 months post Enrollement date N=x	1-2 years post index date N=x	Total collected	Total confirmed not done
NIS	xx	xx	xx	xx	xxx	xxx	xxx
Norfolk	xx	xx	xxx	xx	xx	xx	xxx
KCCQ	xx	xx	xxx	xxx	xxx	xxx	xxx
Nerve conduction	xx	xx	xx	xx	xx	xxx	xxx
Signs and symptoms	xx	xx	xx	xx	xx	xx	xxx
HCP	xx	xx	xx	xx	xx	xx	x
SOI-Hep	xx	xx	x	xx	x	xx	NA
SOI-Card	xx	xx	xx	xx	x	xxx	NA
SOI-Other	xx	xx	xx	xx	x	xx	NA

Table 1: Table showing sample data completeness report layout

Post-Authorization Safety Study (PASS)

A Post-Authorization Safety Study (PASS) is conducted after therapy has been approved by regulatory authorities to gather additional information on its safety or to assess the effectiveness of risk-management measures. Registry-based studies, which utilize data from patient registries, are a common approach in PASS. These studies help identify, characterize, or quantify safety hazards, confirm the safety profile of a medicine, and measure the effectiveness of risk-management strategies. We used data collected in our registry study to analyze and report long term safety and effectiveness of our therapy in real world situations. Our primary objective in this registry study was to assess the safety profile of our therapy under real world conditions which included determining the incidence of selected adverse events of interests related to hepatic, renal events in patients with ATTR amyloidosis patients exposed to our therapy of interest. We were also interested in characterizing the safety of patients in various subpopulations like those with previous liver transplants, or patients with hepatic or renal impairment. The impact of our therapy on pregnancy and infant outcomes was analyzed and reported. Considering data will be collected over 10 years, we have multiple interim analyses planned to provide an insight into the safety and effectiveness of the therapy in real-world situations.

Pharmaco-Economic Analysis

Pharmacoeconomic reports created using registry study data provide valuable insights into the economic impact of healthcare interventions. These reports typically include cost-effectiveness analyses, cost-utility analyses, and cost-benefit analyses, which help in assessing the value of medical treatments and interventions. Data from our registry study was used to track patient outcomes, treatment costs, and quality of life over time. Various economic analyses, such as cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis (CBA). These analyses utilize techniques like multivariate regression modeling and propensity score matching and involve the calculation of key metrics like Incremental Cost-Effectiveness Ratio (ICER). ICER compares the cost-effectiveness of new treatments against the existing standard of care.

Publications and Scientific Conferences:

A journal publication on a long-term observational registry clinical study typically details the methodology and findings of research that tracks patient outcomes over an extended period. These studies often involve linking clinical registry data with electronic health records (EHR) to assess the long-term effectiveness and safety of a therapy. In our registry study, MA team wanted to focus on evaluating the short-term and long-term rates of discontinuation or persistence among patients receiving different treatments for transthyretin amyloidosis (ATTR). Specifically, this article wanted to compare patients who started initial treatment on one of the three classes of approved therapies: RNA interference (RNAi) therapy, transthyretin (TTR) stabilizers and antisense oligonucleotides (ASO).

Sankey plots were employed to visualize treatment patterns from the initial treatment to the treatment at the conclusion of the observation period for the analysis population. Additionally, Sankey plots were utilized to illustrate reasons for treatment switching from one class of therapy to another, including those switches not related to safety or effectiveness. The plots also depicted patients who were on combination therapy and those who discontinued from therapy for various reasons.

CHALLENGES AND OPPORTUNITIES FOR STATISTICAL PROGRAMMERS SUPPORTING MEDICAL AFFAIRS

Statistical programmers in the realm of medical affairs present a unique set of challenges and opportunities. One of the main challenges is handling the vast and complex datasets that come from real-world settings, which often require sophisticated statistical techniques and robust data management skills. Ensuring data

accuracy and consistency while complying with stringent regulatory/publication requirements adds another layer of complexity.

A couple of examples of challenges we faced and how we overcame them are in the following

1. The Medical Affairs Statistical Programming Group (MA-SPG) opted to bypass SDTM mapping and directly utilize source data for analysis datasets and reporting. This decision was driven by the nature of the data received from registry studies, which were derived from secondary use of patient medical records. Mapping this data to SDTM standards posed significant challenges due to compatibility issues, particularly with mapping domains such as exposure. To address these challenges, we retooled several standardized macros and adjusted our programming processes, enabling us to effectively use the raw source data for reporting purposes.
2. Medical Affairs (MA) stakeholders requested that the Medical Affairs Statistical Programming Group (MA-SPG) produce plots that are more visually appealing and concise, essentially "publication ready." In response, our programming team collaborated with the stakeholders to establish standards and developed a suite of macros designed to enhance the visual quality and clarity of the plots, ensuring they meet publication standards.

MA-SPG often faces an unpredictable workload due to the dynamic nature of our work assignments. The demand for real-time data analysis and rapid response to emerging medical queries can lead to fluctuating work intensity. Projects may vary significantly in scope and urgency, with some requiring immediate attention due to regulatory deadlines or urgent publication needs. Additionally, the integration of new data sources and evolving methodologies can introduce unexpected challenges, necessitating quick adaptation and problem-solving skills.

Supporting medical affairs as statistical programmers offers a unique opportunity to enhance our knowledge base and foster professional development. By engaging in a variety of projects, we gain exposure to cutting-edge medical research and real-world data applications, statistical methods and procedures, along with data visualization techniques. This challenging and research-focused environment encourages continuous learning and skill enhancement, as we programmers collaborate with researchers, clinicians, and regulatory experts. Such interactions not only broaden our understanding of different medical and regulatory landscapes but also refine our problem-solving abilities. Additionally, the multidisciplinary nature of medical affairs promotes a holistic approach to data analysis, enabling us to develop a well-rounded expertise that is highly valued in the healthcare industry.

CONCLUSION

Statistical programming plays a pivotal role in the Medical Affairs domain of research-based pharmaceutical companies. It facilitates the rigorous analysis and visualization of clinical data, ensuring that insights derived from research are both accurate and actionable. Through the development and application of advanced statistical methods and tools, statistical programmers enable the effective communication of complex data to diverse stakeholders, including healthcare professionals, regulatory bodies, and patients.

Moreover, statistical programming supports the generation of high-quality evidence that underpins medical publications, regulatory submissions, and health economics and outcomes research (HEOR). By ensuring data integrity and reproducibility, statistical programmers contribute to the credibility and reliability of scientific findings. Their expertise in creating sophisticated visualizations, such as Sankey plots and Kaplan-Meier curves, enhances the understanding of treatment patterns, efficacy, and safety profiles.

In conclusion, the integration of statistical programming within Medical Affairs is essential for advancing scientific knowledge, informing clinical practice, and ultimately improving patient outcomes. As the pharmaceutical landscape continues to evolve, the role of statistical programming will remain integral to the success of Medical Affairs initiatives.

REFERENCES

Anupma Dhanda Farrington, A. G. (2023). The Value and Deliverables of Medical Affairs: Affiliate Perspectives and Future Expectations. *Pharmaceutical Medicine*, 417-424.

E Dawn Flick, H. R. (2023). The Value of Pharmaceutical Industry-Sponsored Patient Registries in Oncology Clinical Research. *The Oncologist*, 657-663.

Heidi Taipale, A. T. (2025). Comparative effectiveness of antipsychotic treatment strategies for relapse prevention in first-episode schizophrenia in Finland: a population-based cohort study. *The Lancet Psychiatry*, 122-130.

Jangili, A. (2024, September 28). *The Role of Statistical Programmers in Medical Affairs Publications and Journals*. Retrieved from linkedin.com: https://www.linkedin.com/pulse/role-statistical-programmers-medical-affairs-journals-jangili-j7afe/?trackingId=%2FA1tqVgiTSiThTUCKeaZvA%3D%3D

Liliana Virzi, P. D. (2023). *The Value and Strategic Implementation of Insights Management.* https://medicalaffairs.org/ (Medical Affairs Professional Society).

Nehad A Abudiyab, A. T. (2022). Visualization Techniques in Healthcare Applications: A Narrative Review. *Cureus*, 2-3.

ACKNOWLEDGMENTS

Authors would like to acknowledge Alnylam Pharmaceuticals for providing the opportunity to work on this paper.

CONTACT INFORMATION

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

Nagadip S Rao
Alnylam Pharmaceuticals, Inc
nrao@alnylam.com