PharmaSUG China 2021 - Paper 70

New Opportunities and Challenges for Statistical Programming Job Position in China Bio-Pharma

Xin Wang, Yu Cheng, Beijing Fosun Pharmaceutical Research and Development Co., Ltd.

ABSTRACT

Back in early 21th Century, China as an emerging market of drug development created new occupations and job opportunities in Asian Pacific and China. Statistical programming - one of these hot occupations in drug development area, is a generic job classification providing career opportunities to talents with statistical analysis capability, data integrity concept, logic thinking and code development strength who use statistical analytical software and perform data analysis in clinical projects. Around 2010, such new job opportunities in China branches of Big MNC pharma cultivated Chinese statistical programming talents of Generation Y capable of supporting FDA, EMA, PMDA, and NMPA filing projects by equipping them with international views and professional technique and soft skills. Now time to 202X, with the progress of China local companies in drug development area and evolving regulatory environment, what are the new opportunities and challenges for Y-ers who have already been the backbone and Z-ers – young talents with infinite potential for next era? To provide some thoughts for talent development in China, we will summarize some general career development needs required and compare some possible opportunities in China Bio-Pharma and that in China Branch of Big MNC Pharma. Opportunities always exist although times are changing.

INTRODUCTION

Career opportunity refers to whether time and environment can provide beneficial support to a person, and different opportunities will arise in different environments. If you want to seize career opportunities, you must consider several points:

1. Career environment
2. Status of industry development
3. Qualification requirements for jobs
4. Connections
5. Potential opportunities and risks

We hope to share some of our experiences and compare the career opportunities of statistical programming positions in the field of clinical development of new drugs in China in the early 21st century with the differences between career opportunities today. We will focus on the comparison between China Bio-Pharma and China branch of big MNC Pharma, although these thoughts might also be applied for companies e.g. CRO or Biotech.

THE DEVELOPMENT OF INDUSTRY AND CAREER ENVIRONMENT FOR STATISTICAL PROGRAMMING IN CHINA

The emergence of statistical programming positions in the field of clinical development of new drugs in China was mainly in the first decade of the 21st century, and gradually developed with global CROs and global big MNC Pharma entering China. The statistical programming position is a branch of the statistical analysis position and is part of the highly differentiated structure of the company's biometrics department. Establishing branches in Shanghai of biometrics organization consisting of data management, statistical programming and biostatisticians formed the earliest headquarters or satellite sites for statistical programming talents in bio-pharma companies in China then. CRO predated pharmaceutical companies entering China. The Shanghai branch was established earlier than Beijing. The purpose of the earliest teams was mainly to hire technical talents to support new drug clinical projects of pharmaceutical companies around the world through the CRO model or the internal CRO model. Since there were very

few senior talents, various companies sent overseas senior leaders to China to lead the establishment of Chinese teams and to cultivate young talents.

At the beginning of the second decade of the 21st century, with the development of China's registration environment, the demand for new drugs in China had emerged. New drugs developed by global big MNC Pharma needed to be approved in China. With the advancement of new drug research and development in China, global big MNC Pharma needed to establish China new drug clinical medicine teams and clinical operation teams to support clinical trials in China, and established investment institutions to cooperate with China R&D teams. Among global big MNC Pharma, the statistical programming team in China had grown and developed, not only due to the advantages of Chinese language and time zone, but also from the increase in clinical research and development needs in China. At the same time, China Biotech established and expanded, and China Bio-Pharma had gradually targeted China's new drug applications, which had contributed to the rapid development of Chinese and global CROs based in China in the early days. In the years after 2010, many new teams were built, especially the branches of global big MNC Pharma and emerging China CROs. The location of the team had also evolved from Shanghai as the main center to two centers in Beijing and Shanghai. Because the duration to conduct clinical trials is usually long, some teams were still in the early stages of their pipelines, and the business demand for statistical programming increased slowly in a short period of time. Although new institutions had been established, the size of the team was relatively stable, and job opportunities in the market had not increased significantly.

In recent years, China's registration environment and R&D environment have developed rapidly. A large number of clinical projects are carried out and are progressing rapidly. A large number of China regulations and guidelines have emerged, for example, the requirements for data submission and data standards are gradually in line with global standards. The demand for a large number of new clinical projects has contributed to the rapid growth of the demand for statistical programming talents in China. The team's location has also expanded from the two major cities of Beijing and Shanghai to other major cities in China. Not only have CRO job opportunities increased, but many self-operated functions of China Bio-Pharma have also been established. The team size expands rapidly. There are so many job opportunities for talents to try, definitely hard to decide which one might be the best fit.

In summary, the development of statistical programming positions in China in the past ten years is rooted in the development of business needs. Global big MNC Pharma have landed in China, and China businesses have risen. The continued stability of this change has led to other changes and developments in statistical programming jobs in China. Stable R&D business needs have driven the statistical programming team to evolve from CRO-based to a combination of internal operations and outsourcing. It has also contributed to the emergence of a large number of job opportunities in Shanghai, Beijing and other cities. The emergence of popular positions has added the value of talents and drove a lot of fresh blood to be in the industry.

At present, statistical programming talents in China include both long-term support for clinical projects registered for approval by FDA, PMDA and EMA, and long-term support for clinical projects registered for NMPA approval. There are R&D centers in Beijing and Shanghai, as well as R&D centers in Nanjing, Wuhan, Xi'an, Shenyang, and Dalian, etc. There are programmers working in self-operated R&D centers of China a Bio-Pharma and global big MNC Pharma, CROs, and Biotech. China's statistical programming industry is still in an upward development stage. It is expected that the development of the industry will bring more new job opportunities and career development opportunities.

THE DEVELOPMENT OF THE INDUSTRY ENVIRONMENT BRINGS NEW JOB OPPORTUNITIES

From the perspective of the development of statistical programming in China, CROs are very adaptable in the new emerging business environment. Many companies develop their business in China and choose to cooperate with the CRO statistical programming team in the early stage. When their business volume stabilizes to a certain scale and their R&D needs increase, some companies choose to establish an internal statistical programming team. The rise of China's new drug clinical statistical programming team has roughly experienced global CRO (global big MNC Pharma outsourcing model), global big MNC Pharma self-operation, China CRO (China biotech and China Bio-Pharma outsourcing model), China biotech and China Bio-Pharma self-operation. At present, the statistical programming team has continued to develop after being established in these different types of enterprises. Compared with job

seekers in the past, the current statistical programming job types are much richer, creating more opportunities, and also adding a certain degree of complexity to job seekers. The environment is changing rapidly, and the needs of enterprises are changing rapidly, and most of the statistical programming positions of enterprises are still in the early start-up stage. What kind of position is most suitable for you? The answer to this question now is much more complicated than it was ten years ago, and the answer is also different from the past. At present, there are many job options, and the complexity of matching your strengths and needs to all various jobs has indeed increased, and as a result turnover rate of statistical programming teams has also risen sharply in the short term.

PROJECTS OPPORTUNITIES BOTH VALUABLE FOR CHINA AND GLOBAL

In the past, due to differences in regulatory requirements, there were significant differences in project requirements in China and in global. Now, the differences in submission requirements, data standards, and study design of China and global projects are gradually narrowing. China projects also have their own unique needs raise. Most of the registration requirements for FDA and EMA have known timelines enabling prepares ahead of deadline for pre-planned requirements. Therefore, the corresponding statistical programming work can be planned in advance and prepared in advance. Both big MNC pharma and big CROs have established systems or processes to increase overall quality and efficiency. The communication mode of China projects seems to be relatively more flexible and mobile requiring statistical programming teams to stand by timely to follow up with regulatory requirements and accumulate submission experience in their daily work. It is difficult to directly use the experience and tools supporting FDA or EMA registration to support fast response request of NMPA, and vice versa. The registration experience of China and global projects is irreplaceable being all valuable to its own business objectives.

CORE TECHNICAL QUALIFICATIONS IN THE NEW ERA FOR STATISTICAL PROGRAMMING

Statistical programming is a project-based technical position. Its core technical ability is to use programming methods for statistical analysis to generate CDISC data and analysis results in line with industry standards. Its core business goal is to ensure high-quality delivery and high efficiency. In the long run, everyone should keep an eye on individual development to advance your project management ability, technical skills and soft skills. Good programming ability and logical thinking are the basic requirements of statistical programmers. Efficient and high-quality completion of statistical programming is the gold standard for measuring excellent statistical programming performance. The correct application of statistical analysis methods and data standards is closely related to the high quality of statistical programming. New technology applications, advanced programming skillsets, standard business processes and full training experience are closely related to the high efficiency. Today's statistical programming work scope and job description set higher bars for statistical programmers, who are expected to be with an in-depth understanding of clinical trials, data standards, statistical analysis and programming techniques. We wish to summarize some mainstream changes and explore the up-to-date important technical qualifications for statistical programmers to think over how to strive for excellence in their jobs:

1. The past data standard does not necessarily to be CDISC standard. The data structure could be either vertical or horizontal. Now for data to be submitted to either FDA, EMA, PMDA or NMPA, it is recommended to submit data compatible to CDISC standards which is vertical for most of all. Statistical programmers are required to be familiar with the concept and application of CDISC and be familiar with the similarities and differences of data standards and analysis standards between versions with different languages and different regulatory needs in different countries.
2. In the past, satisfying the analysis needs of customers only required you to use SAS@ software to generate reports. Current customer needs require statistical programmers not only to use SAS@ to generate traditional reports, but also to be able to combine static reports with visual interfaces. Some companies use R or Python to streamline interactively data review process, and some companies use JMP@, Spotfire@ etc. to leverage the customized interface embedded in these soft wares. The use of SAS@ software has continuously been expanded from BASE and STATS to some new modules in the past ten years. These new demands of development and statistical analysis make our statistical programmers be more proficient in programming logic thinking and more flexible in the use of different analysis tools.
3. The end to end experience of the past statistical programmers started from SAP to the end of CSR (supporting statistical analysis), and the current statistical programmers should start from the

beginning of the plan (Protocol synopsis stage or pre-IND stage depending on study needs) to the end of submission (supporting interactive data display, static reports, statistical analysis, and submission). Now our statistical programmers are required to understand the design being able to foresee various needs instead of being informed. Familiar with data standards and submission standards, be responsible for making decisions on submitting data design, rather than just coding by following orders. Now we are required to be able to handle big data and complex data, and be able to ensure the consistency between Chinese data and English data.

4. Most of the past project milestones could be anticipated in advance, and accordingly statistical programming is relatively paced. You could accept tasks in batches and deliver tasks on deadlines. However, the current regulatory requirements, business requirements, research needs and technical requirements are changing at any time and may be interconnected with each other. The fixed-cycle working model cannot meet the immediate needs of internal medical review or of fast responses to answer questions from regulatory agencies. Everyone's work cycle is shortened, and new risks may appear every day. Current customer needs require statistical programmers to be able to support the delivery of statistical programming in real time. The period from accepting requests to a final delivery has been greatly shortened. The new teamwork model and organizational structure is under exploration.

5. In the traditional teamwork model, when the turnover rate is relatively stable, a paired cooperation model of one programmer closely cooperated with another QC programmer is often used. When the staff turnover rate is high or the project turnaround rate is fast, the cooperation model based on division of labor is often adopted. For paired cooperation model, statistical analysis staff can in-depth study end to end project experience and be responsible for all the work. Though the experience growth cycle might be relatively slow, the project experience might be highly complete. In the division of labor model, the statistical programmers can deeply learn the various changes of the business modules they are responsible for. The experience growth cycle is very fast but might be in pieces. The project experience might be incomplete in the short term. These two models are both very good learning environments for the experience growth of statistical programmers, but each has its own benefits. For paired cooperation model, if the complexity of the project itself is relatively low, the experience growth will be very slow. For the division of labor model, how to divide labor is very important. Long-term fixed division of labor in a certain area may make the project experience become a null accumulation, and the experience cannot be extrapolated. Now, due to more complex business requirements with risks from new issues emerging every day, the paired cooperation model might be more effective than in the past. The division of labor model, because of the time pressure of the project, might set a much higher bar for project managers. When choosing a job, with a clear understanding of the team's business environment and management model, you can target your short-term development needs precisely to your dreaming job.

CORE EXPERIENCE QUALIFICATIONS IN THE NEW ERA FOR STATISTICAL PROGRAMMING

Considering the characteristics of projects for FDA and NMPA submission, we believe that in general, 1) the statistical programming work of FDA submission requires more workload and is relatively more complicated; 2) the infrastructure of most of China Bio-pharma new teams is under developing, there are not many tools available, and might be more challenging when facing risks.

1. If China Bio-pharma new teams supported FDA submission, it might encounter a situation where the workload might increase dramatically with risks still not under control due to the complexity of projects and lack of fitted processes and robust tools. The major risk of statistical programming might be from misunderstanding of the data and inappropriate behavior when programming on data. FDA submissions run by China Bio-pharma new teams, compared to NMPA submissions operated by China branch of global big MNC pharma, require more experienced statistical programming leads onboard and joining projects team before projects move on to their key milestones like NDA submissions. At least when projects are still in early phase build new detailed processes specific for that compound or molecule or indications might be workable for the new team to mitigate potential risks from the challenges in late phases.

2. For China Bio-pharma new teams supported NMPA submission, taking into account the time pressure and finance pressure, it's vital to assign experienced statistical programming leader to simplify the process and make decision when facing changes. More importantly, it's important to equip the team with higher level statistical programming leaders working closely with portfolio leaders to identify potential collaborations or revolutions, or bring innovations to boost fast deliver.
3. To support FDA submissions operated by China branch of global big MNC pharma, which are usually supervised by senior statistical programming leaders in global, the team usually look for qualified statistical programmer to be able to work in a global team who are expected to be well trained and highly sensitive to programming details.
4. For China branch of global big MNC pharma, with mature processes and systems serving specifically for their business, they more focus on recruiting talents with strong implementation capability. Furthermore, they also encouraging leadership development of talents especially young talents looking for innovations from diverse talents.
5. For China Bio-pharma, with numerous opportunities to gradually form at each level its own culture, work environment, work process and standards, they are looking for talents with motivations and capabilities to lead projects and teams in a changing environment. Furthermore, they also are spending a lot of efforts to build up a stable work environment and development area for its young talents to grow up.

OPPORTUNITIES

We try to summarize some of the opportunities that China Bio-Pharma could provide to statistical programmers. We do add some comparison with those in a global big MNC pharma.

YOU CAN CHOOSE DIFFERENT JOB ROLES

1. Statistical programming role, mainly engage in the analysis of statistical methods and analysis data.
2. Data programming role, mainly engage in data collection, transformation and tabulation.
3. Technology role, mainly engage in the technical work like tool development, system construction, and code optimization.
4. Outsourcing oversight role, mainly engage in vendor and project management.

THERE ARE MANY SENIOR TALENTS AND A GOOD LEARNING ATMOSPHERE

1. Rich experience in China and global submissions and standards like CDISC, etc.
2. Diverse programming experience in using SAS@, R or Python.
3. Several groups working on Chinese translations of guidelines.
4. Working together for domestic and global submissions readiness.
5. Well trained leaders good at dealing with unexpected incidents from projects and from the team.
6. Cooperative communication and learning are barrier-free when speaking Chinese.
7. Eager to learn and share to solve problems together.
8. Collaborate with global experts in areas like CDISC initiatives, etc.

MORE OPPORTUNITIES TO TAKE LEAD

1. Project leads are highly motivated seeking high efficient management experience as not much available process or tools to use directly.
2. The license-in project and the license-out project provides opportunities to improve your soft skills e.g. cooperation and communication in extreme complex collaboration model.

3. Therapeutic areas e.g. oncology or vaccine provides talents with long-term in-depth leading opportunities in TA standards and TA tools.
4. The needs of using both Chinese and English standards adds its complexity in standards implementation.
5. System development area look for more leaders.

CHALLENGES

SAS@/STATISTICAL AND MEDICAL KNOWLEDGE

To beginners (1-3 years' experience), with dozens of reference studies and usually a central location of system macros which are called for any standard code for deviation to generate standard Tables, Figures and Listings (TFLs), and even an internal GUI (graphical user interface) to drag and drop or point and click their way to create standard TFLs, SAS@ programmers are able to take over their jobs quickly at the beginning while working in a big MNC pharma, though one may not know every detail how the variables are coded or what the detailed algorithm of each step which is due to the mature system and documentations especially when the studies taken are designed similarly. Hence, this is a challenge to develop your own style of SAS@ programming, and to learn the statistical method and medical knowledge at the early stage of your career.

While working in a China bio-pharma, as there may not be a set of full-developed reusable macros for use, for one particular domain in one particular study, you have to write your own codes, complete from scratch the needed meta files, find and understand the application of the CDISC standards with a data collecting system designed relatively weak in standardization, or may be involved in the development of macros, which push you to ask questions and learn in order to fulfill your task. It will be really tough for junior programmers to accomplish all of these at the same time. And there is a great demand to develop and validate study meta files, data specifications, and standard codes or system macros and to implement training programs in China bio-pharma.

CDISC STANDARDS/SUBMISSION RULES

In China branches of big MNC pharma, programmers can modify the submission datasets based on existing integrated data standards, and prepare the submission documents using a similar experience as is submitted to FDA/PMDA/EMA with only translation needed at the last step, which is usually not the responsibility of programmers.

While this is another big challenge for China bio-pharma, with less experiences in submission and growing experience communicating with CDE (NMPA's Center for Drug Evaluation) about data standards, and moreover, the controlled terminology in Chinese characters may not be standardized enough, although the root cause may lie in the diversity of source data collecting rules. For programmers, the experience towards submission can only be gained according to Pinnacle 21 validation rules when producing the submission documents (SDRG, ADRG and Define.xml) which is usually not real-time answering the questions from CDE. As for the implementation of CDISC standard in China bio-pharma, as data standard are mainly developed per particular study, the data quality and integrity may be highly impacted by the programmers' experience as well as the knowledge and comprehension to SDTMIG or ADAMIG, and Pinnacle 21 validation rules, etc.

TIME MANAGEMENT/LEADERSHIP ROLES AND RESPONSIBILITY

Despite the milestones and deadlines that you set, to schedule or complete a project can be difficult for a relatively junior programmer, because the timeframe for programming is often uncertain, much of which is decided by the study team in line with data collection timeframe which is always changing or might be shortened suddenly by a new department strategy or an urgent regulatory request. Iterations will likely need to be made before the experience being gained enough. Eventually you will have a better sense for scheduling a project, although we believe there may be few differences between big MNC pharma and China bio-pharma in programming aspects. Moreover, the communication with study team members can

be challenging but in a slightly differential way, with programmers feeling hard to communicate to abroad counterparts in a big MNC pharma or unfamiliar counterparts in a China bio-pharma.

Next, you will have an important decision to make:

Do you want to stay on the current career path?

Or, do you want to develop your technical skills or useful programming tools?

Or, do you want to be in charge of projects, or lead a team of your own?

Taking a leadership role in your organization is one of the most challenging tasks you will face, because at this point, you are no longer accountable for only yourself anymore. You will also be responsible for other people as well and will need to ensure that the projects are finished by deadline without risks – all programming rules are aligned with team members' decision. And your team will look to you for guidance along every step of the way. Relatively, there is few room for errors as a project leader in big MNC pharma with few opportunities here, whereas in local pharma, taking this role might be easier but with more uncertainties.

CONCLUSION

China Bio-pharma provides talents with several jobs and development opportunities. Although the environment is still challenging full with risks, there have been a lot of peers working together for within-organization solutions or external collaborations.

CONTACT INFORMATION

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

Xin Wang
Beijing Fosun Pharmaceutical Research and Development Co., Ltd.
wangxin9@fosunpharma.com

Yu Cheng
Beijing Fosun Pharmaceutical Research and Development Co., Ltd.
chengyu@fosunpharma.com