Paper 125-31

Moving from Programmer to Statistician – A Statistical Mentoring Program

Emily Anne James, Biogen Idec, Cambridge, MA

ABSTRACT

Programmers and Statisticians typically interact on a daily basis, thus making the relationship between the two fundamental to a successful project. Did you ever wonder what it would be like to be in the shoes of a Statistician? I'm sure most of us programmers have wondered what it would be like but how many actually have the desire to try it out? This paper will tell the story of a Programmer's experience working as a statistician.

INTRODUCTION

I have always had a strong interest in statistics. With a bachelor's degree in sociology, statistics and methods were my preferred and most rewarding courses. I attended graduate school for a Master's in Public Health in Epidemiology & Biostatistics. After graduating, I began my career in the biotechnology industry. I was hired as a Statistical QA Analyst, mainly to check statistical output and analysis datasets and parallel program data points. I viewed this position as a stepping-stone to something more functional.   I was in the Statistical QA role for one year and then moved into a Statistical Programming position. Programming was much more stimulating for me yet I missed utilizing my statistical skills. I wanted to take on more of a statistical role to determine if that is the direction I want to go in the future. I worked side by side with Statisticians and had a general sense of what they did but I wanted to observe their activities on a daily basis.

RIGHT PLACE AT THE RIGHT TIME

After expressing this interest with management in statistics, a plan was developed. My goal was to obtain hands-on experience with a variety of tasks statisticians perform. At the same time there was a need for more statistical resources in the drug program in which I had been working. The statisticians were overloaded with project work for the upcoming quarters. In addition, the Lead Statisticians were anxious to get more mentoring experience. This was a perfect opportunity for someone like me who had the statistical background and interest to provide statistical support while learning more about the role of a statistician in the company. My extensive experience in the drug program coupled with my statistical skills allowed me to pursue this opportunity. Cleary, I was at the right place at the right time.

I drafted a proposal for a Statistical Mentoring Program with the goal of developing a solid understanding of the daily activities and responsibilities of Statisticians. This entailed using my statistical knowledge and skills to provide support to Lead Statisticians on various projects. After discussion with management, it was decided that I would split my time 50:50 between programming and statistics as needed to meet project timelines. The programming projects were kept separate from the statistical projects so that I could maintain objectivity.

STATISTICAL TASKS

The majority of my statistical work was focused on preparing for an interim clinical study report for a large, complex study. There were tight timelines to submit the interim clinical study report with minimum resources. My statistical tasks ranged from data checking; to creating specifications for analysis datasets, tables, and figures; to parallel programming.

The statistical tasks included but were not limited to:
- Drafting table shells
- Creating specifications for analysis datasets
- Creating specifications for tables, figures, and listings
- Identifying first look analyses
- Performing data checking
- Parallel programming analysis datasets
- Checking data results and parallel programming tables, figures, and listings

The Lead Statistician acted as a mentor by delegating statistical tasks, answering questions regarding assignments and standards, reviewing my work, and providing feedback. When the proposal was drafted, there were many projects potentially in need of help.

I did a lot of the preparation work for the Statistical Analysis Plan, creating specifications for analysis datasets including efficacy and pharmacodynamics. I drafted table shells and wrote specifications for all of the tables and figures. I created data checking programs to ensure the data was accurate and consistent and communicated and followed-up with Data Management team members to resolve data issues. For the programming and analysis stage, I was the statistician for the baseline and safety sections. I parallel programmed analysis datasets, baseline tables and safety tables and listings.

PROGRAMMING TASKS
Additionally, I acted as Lead Programmer for the study. I delegated work to various programmers, briefed them on the nature of the study, assigned parallel programmers for QC of analysis datasets, communicated with programming management regarding resources, followed up with programmers to ensure we were on target to meet timelines, tracked programming and QC progress, and fielded questions from programmers regarding study issues, timelines, data, and statistical specifications. I extracted raw data periodically from the clinical data warehouse, kept team members informed regarding progress and reminders, held study update meetings, and ensured data issues were being tracked and presented to Data Management. These responsibilities as the Lead Programmer could overlap with those of the statistician, depending on where and with whom you are working.

A DAY IN THE LIFE OF
The typical day in the life of a junior statistician varied depending on the stage of the analysis. After thorough review of the study protocol and draft analysis plan, my day was spent drafting table shells and writing specifications for analysis datasets and tables, figures, and unique listings. All of my work was sent to my statistical mentor for review and I made modifications where necessary. Once the programming began, I was parallel programming and checking analysis datasets, answering questions from programmers regarding the specifications, parallel programming tables and figures, and checking results for consistency and accuracy. After the programs and analyses were finalized, the study report drafts were circulated for review by the study management team. Because of the scale and intricacy of the study, the review generated various changes and additional analyses. Additional analyses and modifications triggered another cycle of writing specifications, delegating tasks to programmers, and parallel programming and checking the output. Once the study report was final, I moved some of the tables, figures, and listings into the final report document.

COMPARISON OF PROGRAMMING AND STATISTCAL TASKS
There were a lot of advantages to working as a Statistician with my programming background. My familiarity with the drug program, programming techniques, and conventions allowed for a smooth transition. I already had the programmer perspective from my experience programming similar analyses. This was extremely useful while developing specifications for analysis datasets, tables and figures. I would think what type of information would I need as a programmer to be able to code this correctly.   One of the features I particularly liked about the Statistician role was the variety of tasks involved, not merely coding all day as a Programmer is likely to do. As a statistician on any given day, I could be doing anything from creating specifications to parallel programming to checking data and results.

I discovered that there were quite a few drawbacks to acting as a Statistician that I wasn't used to dealing with as a Programmer. The medical team made many changes to the study report, which involved revising table shells, specifications, and programs after the fact. There were endless interruptions from Programmers asking questions regarding specifications, timelines, and study issues. This could have been due to the fact that there were many temporary programmers working on the study. Directing the Programmers was not always easy. Acting as the Statistician for the baseline, safety, and listing sections involved working with Programmers and QA Analysts to answer questions and explain study issues and specifications. There was a tremendous responsibility on my part for this study being the Lead Programmer who could not do any of the programming since I was also the Statistician who developed many of the specifications. I was responsible for making sure the programming was complete and correct in a timely manner, yet I could not participate as a Programmer to ensure this. I truly experienced what it was like to be in the shoes of the Statistician. The Statistician is responsible for the work and waiting for the Programming deliverables.

The biggest concern I found working on this study as a Statistician was discovering the considerable gap in skill level among Programmers working on the study. I held high expectations of all of the Programmers since the main Programmers on the study were so proficient. It was very difficult for me to remain calm and content when the programming was not being done correctly or in a timely manner. This was only the case when we assigned a few inexperienced programmers to handle some of the tables. With the tight deadlines, I expected everyone to pull their weight to meet the timelines. Working as a Programmer, I had never experienced the burden of accountability that the Statistician holds.

CONCLUSION

My experience in the statistical mentoring program was extremely valuable. It gave me tremendous insight into a Statisticians' point of view and daily stressors. I was able to develop my statistical skills and knowledge in the industry by acting as a Statistician for baseline and safety sections of a clinical study report for a pivotal study. After this experience I was confident that becoming a Statistician was not the career move I wanted to make. I would not have been able to make this decision without this experience with the mentoring program. I am grateful to my company and department who allowed me to fulfill this opportunity while maintaining my programming position. This preparation permits me to take on a greater variety of tasks at work as a Programmer with statistical projects. This was an excellent experience overall. Not only was it challenging and enlightening, it helped me make a very important career decision. I believe this has made me a better Programmer overall.

CONTACT INFORMATION

Contact the author at:
      Emily Anne James
      Biogen Idec
      14 Cambridge Center
      Cambridge, MA 02142
      Phone: (617) 679-3825
      E-mail: Emily.Anne.James@biogenidec.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.
Other brand and product names are trademarks of their respective companies.