EVOLVING ROLE OF STATISTICAL SAS® PROGRAMMERS IN THE PHARMACEUTICAL INDUSTRY

Venkata Sekhar Bhamidipati, Merck & Co., Inc., North Wales, PA
Haiping Zhou, Merck & Co., Inc., North Wales, PA

Abstract

The high level of competition in pharmaceutical industry has recently led to mergers, organizational changes within companies, and other various types of cost-cutting measures. This trend has significantly impacted the role of statistical SAS programmers. To catch up with the fast-paced changes, a new or experienced statistical SAS programmer has to expand their skill set to meet the new expectations of the industry. Statistical SAS programmers no longer can contribute effectively with only SAS knowledge. In addition to SAS knowledge, statistical SAS programmers must have a good grasp of clinical trial database setup, eCRF/CRF development, the data collection process, etc. Most importantly, statistical SAS programmers need to hone their communication skills, in order to be able to interact effectively with different stakeholders in the new working environment.

This paper will examine the traditional role of a statistical SAS programmer in the pharmaceutical industry and contrast it with current expectations, and review the new skills required to successfully perform the job. It will also attempt to provide perspective to new aspiring statistical SAS programmers and experienced statistical SAS programmers.

This paper will also delve into a few tools and techniques which a programmer may find helpful when performing the “new” statistical SAS programmer role.

Introduction

The role of a statistical SAS programmer has been vastly expanded due to cost-cutting measures, limited resources, and outsourcing strategies implemented by many companies in the pharmaceutical industry. The current expectations of statistical SAS programmers are significantly different from those of traditional statistical SAS programmers. They are asked to participate in a wider variety of activities, such as clinical trial database setup, CRF annotations, medical monitoring, data quality checking, management of outsourcing partners, etc.

This paper will focus on some of the new activities in which current statistical SAS programmers are expected to participate. It will also discuss some of the tools and techniques that may be utilized to enhance job performance.

Traditional Role of a Statistical SAS Programmer in the Pharmaceutical Industry

The traditional expectation of a statistical SAS programmer is to work with a 'clean' database to create analysis datasets, and produce SAS based reports as guided by the protocol and statistical analysis plan (SAP) provided by the statistician.
The following flow chart shows the traditional role of a statistical SAS programmer.

![Flow chart]

**Current Expectations of a Statistical SAS Programmer in the Pharmaceutical Industry**

The current expectation of a typical statistical SAS programmer is to participate in any or all of the clinical data management activities, which include clinical trial database setup, CRF/eCRF design, annotations, mapping, medical monitoring, and data quality checking. He/she is also often expected to manage or work with outsourcing partners, in addition to the traditional role responsibilities.

The following flow chart shows the current expectations of a statistical SAS programmer.

![Flow chart]
Clinical Trial Database Setup

Statistical SAS programmers are often expected to participate in clinical trial setup in order to elevate any programming issues up-front. They are often expected to participate in suggesting the addition or removal of edit checks in data collection tools. In order to provide constructive suggestions, they should have a good knowledge of current data collection tools like Inform, ClinTrial etc.

They are also often asked to participate in clinical trial data(metadata) setup like phasing, treatment groups, baseline flags etc., which plays a critical role in the formation of database structures. Statistical SAS programmers involvement in the metadata setup is often highly valued because a programming-friendly database structure can save lots of programming manpower.

CRF/e CRF Design/Mapping/ Annotations

CRF/eCRF design is another critical element for the success of a clinical trial. A good CRF/eCRF design can reduce the confusion during the data collection which ensure the accuracy of reporting in the end. Since statistical SAS programmers have an in-depth knowledge of each data point utilized in the programming, their involvement can help to ensure that critical data points are collected efficiently via CRF/eCRF. Take concomitant therapy reporting for an example, in order to differentiate a prior therapy from a concomitant therapy, either a date or a flag needs to be collected via CRF/eCRF. Without that, it will be extremely difficult to differentiate prior therapies from concomitant therapies in the programming.

In addition to CRF/eCRF design, statistical SAS programmers are also often expected to play a significant role in mapping each and every CRF/eCRF data point in to the database. To perform this activity, they should be familiar with industry quality standards, guidelines and procedures. For example, a good knowledge of the Study Data Tabulation Model (SDTM), which defines a standard structure for study data tabulations that are to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA), is definitely helpful.

Statistical SAS programmers also often expected to participate in annotation of all the database domain variables in CRF document, annotated CRF document to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA). Abode Acrobat professional is widely used for this. A good knowledge of this tool is helpful to perform the job effectively.

Database User Acceptance Testing (UAT).

A good database is critical for final reporting. Statistical SAS programmers are often expected to be involved in the database UAT because of the close relationship between database quality and accurate reporting. Database UAT is done after the database design using mock data to check each and every data point is collected properly. A key data point missing or incorrect collection can have a negative effect on lots of other seemingly-unrelated data points and also negative effect the programming. For example, a missing study medication date can effect phasing. A good statistical SAS programmer can definitely contribute a lot in the establishment of a database in high quality.

Medical Monitoring, Data Quality Checking

Statistical SAS programmers are often expected to support data cleaning and critique actual values. A good statistical SAS programmer is capable of looking at the data as a whole, not as an individual data points or forms unrelated to each other, to detect incorrect data collection patterns or safety-related trends in the trial. This will provide valuable input to the responsible groups like
clinical, statistics etc., so they can adjust the data collection guidelines or trial development
directions in a timely manner.

Manage/Work with Outsourcing Partners

Due to cost-cutting measures, limited resources and various collaborations, statistical SAS
programmers are often expected to manage or work with various types of groups like clinical,
statistics and programming etc., outside their own companies. Some times they are also
expected to work as project lead not just as programming lead. Since each company has its own
culture, they have to be capable of adjusting themselves, as needed, so as to create a
harmonious working environment where people with different backgrounds and different locations
can work together on the same project, following the standards defined by the sponsor's
company.


To be able to accomplish these new tasks, excellent communication skills, coupled with a firm
knowledge of various types of tools, like Microsoft Office suite, Abode Acrobat Professional, R,
S-Plus Application, Oracle functionalities (views, sub-queries, functions…etc) in a reporting
environment, databases such as Oracle and SQL Server, data capture tools such as InForm and
ClinTrial and data review tools such as Integrated Review (IReview, JReview), have becoming
increasingly important for statistical SAS programmers to obtain. Each tool is used in one or more
of the areas discussed in this paper.

Communication

Statistical SAS programming is no longer a heads-down coding activity. As a statistical SAS
programmer, he/she has to interact with a variety of groups (Clinical, Statistics, Data
management, Regulatory, vendors, and outsourcing partners, etc.) to gather requirements and
specifications or to coordinate activities. Therefore, excellent verbal and written communication
skills have become extremely important, especially true if team is global. A poor communication
will result in confusion and low productivity.

Strong planning and organization skills are also requested by many hiring companies. Since a
statistical SAS programmer can be asked to be involved in several activities simultaneously, it is
essential for him/her to be able to multitask with good prioritization. It can be very difficult to
survive in the current world if you cannot multitask.

Tools

The ability to effectively use various types of existing software applications is equally important.
A good example of this kind is the wide use of Abode Acrobat in CRF/eCRF annotation. Another
example is the I-Review tool commonly used by current pharmaceutical companies. This tool can
be used to review live data. It can also be used to analyze clinical data from clinical database
management system. Statistical SAS programmers may be required to program various types of
reports using this tool. Therefore, it is very helpful for them to be familiar with the structure of this
tool and type of programming language used in this tool. There are many ways to do quick charts
and plots and summary statistics, besides SAS, like S-plus, R, Microsoft Power Point and Excel.
All these tools can help statistical SAS programmers to excel in their new job roles.

Conclusion

In order to be competitive in the current pharmaceutical industry, new or experienced statistical
SAS programmers can no longer rely on SAS knowledge only. They have to train themselves to
be ready for many new tasks beyond their traditional roles.
ACKNOWLEDGEMENTS:

We thank Rich Lowry, manager scientific programming, Merck & Co., Inc. for his cooperation and collaboration in reviewing this paper.

REFERENCES:

SAS Programming in the pharmaceutical industry, By Jack Shostak. All rights reserved


CONTACT INFORMATION:

Your comments and suggestions are valued and encouraged.

Contact the authors at:
Venkata Sekhar Bhamidipati  Haiping, Zhou
Merck & Co., Inc. Merck & Co., Inc.
Scientific Programming Analyst Scientific Programming Analyst
Email: venkata_bhamidipati@merck.com Email: haiping_zhou@merck.com
Phone: (267) 305-1632 Phone: (267) 305-8231