

Life as a new programmer

Jeremy Gallien, Novartis, Basel, Switzerland

The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Novartis. Novartis does not guarantee the accuracy or reliability of the information provided herein.

ABSTRACT

Starting as a programmer in the pharmaceutical industry can be challenging. Whether it is your first job, you are a trainee or you come from another field; the difficulties are numerous and varied.

This can lead to a lot of frustration, nervousness, and make your start a nightmare. I started work as a trainee, and I have now taken a look back at my time as a novice statistical programmer. I have summarized what problems I encountered and how I solved them.

The aim of this paper is to help new starters in order to facilitate their beginning as a statistical programmer, to give tips and tricks and highlight common mistakes. To achieve this, I will present scenarios to illustrate problems encountered in the everyday life of a beginner.

INTRODUCTION

The aim of this paper is to help the beginner. In a first section, I will briefly describe the work of a statistical programmer. Then, I will explain some situations that statistical programmers encounter every day with the do's and don'ts. Finally, I will give some tips and tricks to improve as a statistical programmer

WHAT IS A STATISTICAL PROGRAMMER?

In simple terms, a statistical programmer creates datasets, tables, listings and figures, according to specifications describe in the protocol and Statistical Analysis Plan, using data stored in a clinical database.

In order to do this, the statistical programmer has to work closely with the statistician and the data manager to provide input for the data quality and analysis needed.

As a new statistical programmer, you will have to learn lots of things. Data can be varied depending on what Phase or what therapeutic area you work in.

A DIFFICULT BEGINNING

When I started as a trainee, I experienced many difficulties. The pharmaceutical industry has its own vocabulary, acronyms, complex data (for example, laboratory data, RECIST etc.). Additionally, you have to interact with lot of different functions within a project.

Below I will explain some situations that I have encountered.

DURING THE PROGRAMMING PHASE OR THE QUALITY CHECK

As a statistical programmer, the quality, traceability and reproducibility are very important. All activities need to be well documented and controlled. This is why almost all datasets and outputs have to follow a validation process. It can be done with a double programming, or a code review. You can also refer to the Good Clinical practice (GCP) for more information.

PhUSE 2014

You might experience few situations during the validation process. As a new starter, you will certainly make errors, and this is normal. Therefore, do not rush! Take your time to really understand the data, the specifications and the programs carefully, check all the steps in your program. This will help you in the mid and long term as you will become familiar with the process.

If you cannot match with the developer, do not tell yourself "I'm a junior without experience, I must be wrong". Check a few cases manually, do not hesitate to make screen capture from your screen, create intermediary dataset/datasetsteps, select a patient or a group of patients, select variables of interest, read the specifications and the SAP one more time to make sure that you implemented the correct logic, and if you still think you are right after that, discuss with the developer, even if she or he is more experienced. I am sure she or he will be happy to discuss with you! Even if at the end you are wrong, you'll have learned how to implement the correct logic for your next assignment.

Asking questions reflects well on you because it shows that you are showing interest about what you are doing. Do not fall into the trap of programming something that is in the specification document without using critical thinking. If you disagree, then always say so, and provide your rationale.

When you create an output, consider the following (for instance):

- Do the numbers in the output make sense? (e.g.: Can this patient be really 256 years old? Is 3.22 meters a reasonable measurement for person's height?)
- Are the numbers in the table consistent across other outputs? (e.g. : the number of deaths in Table X should match the number of death events in Table Y)

If you ask yourself such questions when you create a report, you will prevent a lot of issues during the validation and you will save a lot of time in the later stages of the project.

COMMUNICATION

I am sure that, very early on in your new job, you will experience meetings and teleconferences with your colleagues. My suggestion here is to prepare a short introduction about yourself, which should last no longer than one or two minutes. Doing this before the meeting will help you reducing the stress due to meeting your colleagues for the first time.

What you will learn very soon from this type of meeting or while reading e-mails, is that the world of the pharmaceutical industry is full of abbreviations! You can be lost after only few minutes. My suggestion here is to ask your mentor or a senior colleague to provide you with the list of the acronyms and to speak with them after the meeting if there is something you did not understand.

Work closely with your statistician. If you do not understand the specifications, that may be because they are not clear, do not hesitate to ask questions to the statistician. You can also make suggestions on how to clarify or improve the descriptions of the analyses and of the data-handling in the SAP.

SKILLS AND TOOLS

As a new programmer, you will face few situations where you do not know how to perform an analysis. This can be a programming skill, or more about the concept. In this case, you should spend a bit of time to search by yourself, but not too much (no more than two hours). If after that, you still did not find a solution, then ask a senior programmer. But always try to find the solution by yourself first, you will learn a lot doing so.

Sometimes, you will need a specific tool to perform a task. Often, this tool already exists but you are not aware of its existence. Instead of spending time developing, ask others programmers if they know of an existing tool that could help you.

TIPS AND TRICKS

DOCUMENTATION

You should always have the protocol, the CRF and the SAP at hand. These documents are not only helpful, they are at the core of what you do. If you have any doubts about something, then just check these documents.

Also you should be aware about the International Conference on Harmonization. This group is working in the United States, Europe and Japan to develop common regulatory guidance. You must know the 21 CFR-part 11 because any work you will submit to the FDA must follow this guidance. Standard Operating Procedure (SOP) should follow this guide, but you can also learn by yourself.

IMPROVE YOUR PROGRAMMING SKILLS

At the beginning, the best thing to improve your programming is to look at other's programs. You will learn about procedures that are the most commonly used in the pharmaceutical industry. Also, your company is likely to provide trainings to its employees, either about programming or statistics. Be sure to seize this opportunity to deepen your skills!

You will develop your own style, take time to learn things on internet to improve your programming skills. If you need to use an existing program, don't only run it and check the log. Ensure you understand each steps in the program to check if it is doing the good result.

PhUSE 2014

KNOW YOUR STUDY

In my opinion, a good statistical programmer must know the study. Each disease/drug has their own effects. To do this, you can start by reading the protocol. The protocol is something akin to the “operating manual” to ensure that everybody is in the same page.

Then, you can read the SAP. This document presents all the analyses needed for the study, and how the data will be handled. Read it carefully, do not hesitate to ask the statistician if you do not understand a part, you can also make notes and highlight important part. Make sure you are always using the latest version.

Don't take specification as 100% accurate! As explained in the first part, ask questions about derivation. If a derivation doesn't make sense, make sure you question it.

LEARN ABOUT STATISTICS

Statistical programmers are coming from many fields. And some never used or even learned about statistics before. So if you can participate in a training, or even read books or article by yourself, few statistics that are commonly use in the pharmaceutical industry (here I think about the analysis of variance (ANOVA) and the survival analysis), it will help you in the future. Knowing about statistics will help you determine if the results make sense or not. You do not need to be an expert, but a bit of knowledge is really appreciable.

FOLLOW THE GUIDANCE PROVIDED TO YOU

Take the time to clearly understand the SOPs. They will show you how a statistical should work in, and like this you can support your team to highlight any times we are deviating from process.

Don't hesitate to attend training, even if they are not mandatory. It can be for programming, regulatory affairs or communication. This is a good way to improve your skills!

CONCLUSION

Starting to work as a statistical programmer is a challenge but it is also exciting. I hope this paper will help you on the path to become a senior statistical programmer and that it will make your start easier.

CONTACT INFORMATION

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

Jeremy Gallien
Statistical programmer
Novartis Pharma AG
Fabrikstrasse
CH-4002 Basel
SWITZERLAND
Work Phone: +41 61 696 41 60
Email: jeremy.gallien@novartis.com
Web: www.novartis.com

Brand and product names are trademarks of their respective companies.