

Paper PM02

Onboarding a junior statistical programmer in 10 weeks: My Story From University to First Analysis Delivery

Pieter Coppens, SGS Life Sciences – Clinical research, Mechelen, Belgium

Maarten Borremans, SGS Life Sciences – Clinical research, Mechelen, Belgium

Annelies Van Zeveren, SGS Life Sciences – Clinical research, Mechelen, Belgium

ABSTRACT

Young graduates with a background in research, statistics or life sciences are suitable candidates for a position in (statistical) clinical research. They are typically willing to learn, immediately available and provide fresh opinions. However, a lack of experience in clinical research and/or practical skills limits their immediate employability. Using a specific onboarding process, it is possible for such young graduates to complete a first study delivery within 10 weeks after their initial working day.

The paper describes the onboarding from the point of view of a new employee (the first author). It starts with a description of all aspects of this process, as well as the resources and people involved. Further, the timeline from the drafting of an individual training plan, over the onboarding itself, to the eventual delivery of the first analysis will be discussed. Finally, a brief overview of the evaluation process and a reflection conclude the paper.

INTRODUCTION

Clinical research is a field with very high demands in terms of standardization, regulation and, of course, quality. As such, it can be a challenge for young graduates to become fully operational employees in this environment. However, by using a tailored, gradual onboarding process, it was possible for me as exactly such a young graduate to learn the necessary skills and send out a first analysis only 10 weeks after my first working day.

This paper will present an overview of all the steps leading up to this first delivery as told from my personal experience. Before presenting a chronological overview, an overview of the people involved as well as of the contents of this process is presented. At the end, the evaluation process is briefly discussed, followed by my own reflections on my experience.

Please note that as this is an account of my own onboarding experience, it is more a case study than it is a description (let alone analysis) of the general onboarding policy at the statistics department of SGS Life Sciences. So as this is a story told from my perspective, some personal background before diving into the subject matter itself may be of use to the reader.

I obtained a bachelor in electronics and mechanical engineering in 2008, followed by a master of biomedical engineering two years later. After a year studying mathematical engineering combined with a teaching role, I started a doctoral program at the faculty of science about engineering education, defending my PhD dissertation at the end of 2016. Although I did not have a strong background in statistics, I had some knowledge on the subject from my master and practical experience from my PhD. I had even less experience with pharmacology, but I did have a (very) basic understanding of biomedicine from my master degree. Research and programming was something I did gain a lot of experience with during my PhD. The story below starts in April 2017, when I started as junior Biostatistical Analyst (a position similar to that of a statistical programmer) at the statistical department of SGS.

PEOPLE

My onboarding process was not only a matter of me reading through required documentation and watching online videos on my own (although that was also part of it), but it involved many different people. The first point of contact was my manager, who is the overall responsible for the onboarding. In addition to organizing the onboarding process, the manager also ensures a regular follow-up, adjusts the schedule where necessary and provides part of the in-company training. A second important person, especially at the beginning, is the departmental secretary. She took care of administration, but also of plenty of small practical necessities.

Then, there are the direct colleagues. This includes a dedicated coach, who is an experienced colleague and is there to answer my more complicated questions, to help with programming and also to give feedback. Other colleagues, especially those in the immediate vicinity of my desk, also helped with smaller issues or practicalities. These colleagues are not only the more experienced ones, but also some who started a couple of weeks or months before me. Those less experienced colleagues often have a fresher memory of practical or administrative requirements that more experienced colleagues only dealt with longer ago, and could often understand questions quicker than more experienced people because they struggled with similar issues recently.

A final group of people involved in the onboarding process are external trainers. This includes external companies (e.g. SAS/STAT® Institute) as well as in-company staff that specializes in training, both inside and outside the statistics department.

All these people played an important role in my development, either implicitly or explicitly. Of course, scheduled training sessions are where everything gets started, but it was (and is) even more important to know which people to reach out to for a specific issue or question I had/have. Although that was not always clear at the beginning due to the number of people involved in the onboarding process, it is still common for me and other colleagues (both junior and senior) to ask each other for help, be it for a specific question about our in-house programming macro's or a general question on how to best tackle a complicated programming task. Finally, it is of note that virtually all of these people work office-based, making it easy and fast to interact with them on a daily basis.

CONTENT

I had to learn a myriad of new topics, ranging from how to refill the coffee machine over custom programming knowledge all the way to ICH GCP requirements. Most of these topics can be classified into 3 broad categories: practicalities, background information and daily work.

PRACTICALITIES

Especially at the beginning, the relatively mundane aspects of any new job are an important aspect of a good onboarding process. To me, it was very useful that things such as building access, office supplies and where to find coffee were taken care of and explained on the first day. This made it possible to focus on other tasks that are more part of the job description. Even though this aspect will not be discussed further in this paper because it is similar for virtually all new employees across multiple industries, it is still mentioned because practical issues can absorb a lot of time and energy when not taken care of. This is especially true for a new employee, for whom everything takes more time and energy as it is all new.

BACKGROUND

This category includes training and knowledge of topics that are not immediately relevant for my daily work, but are very much necessary to understand how that daily work fits in 'the bigger picture'. Topics in this category made it possible to make better decisions even during programming and to interact with other partners such as sponsors or other departments in a more meaningful way.

A first topic in this category is learning about the structure of the company. This ranges from very high-level information with very little if any impact on daily work (e.g. that there are also mining, agriculture and food, energy etc. divisions in the company), over how other departments within life sciences and biometrics function (some impact in interactions with those departments) down to the structure of the statistics department (important impact on daily work).

The second topic in this category are the regulatory requirements, such as ICH-GCP and the declaration of Helsinki. It also includes training of general company policies regarding ethics and safety. These standards serve as a backbone to the research I am a part of as an analyst, and are in the back of my mind when discussing how to present a certain topic, when writing documentation, etc. It also explains a lot of why clinical research is done a specific way.

Finally, there is information about clinical trials on a more detailed level. This includes examples of trial designs (the difference between phase I-IV trials, parallel versus cross-over design, etc.), but also how different departments are involved in a trial. These topics start having a more direct impact on the daily work. An example is who to talk to about issues with the database and why it may take a while for an issue to be resolved (data management sending a query to the site, an external vendor who has to send an update ...). This also made it easier to understand what happens with an analysis once it is finished, which in turn helps when setting up the analysis as I could keep in mind how it will be used later on.

Most of the above did not have a big influence on how my first project was delivered. However, these topics do play a bigger role in subsequent projects and other in-company tasks, as well as during the interaction with other departments and with clients.

DAILY WORK

Arguably the most important category, especially while working on my first project.

This category contains first of all an introduction to the standard operating procedures (SOPs) and working instructions of the statistics department. These specify the day-to-day functioning of the departments and provide guidelines on how an analysis is to be done. An example is the quality control: this is done when a first version of the analysis is finished by one person, locked for editing and then the results are verified by an independent second person, who provides feedback. Needless to say, I was person one in the first project. Another example is learning how documents such as the statistical analysis plan (SAP) are structured, updated and used during an analysis.

Secondly, this includes trainings about the structure of SDTM and ADaM datasets. The latter not only requires insight in the CDISC standard itself, but also very specialized training on the in-house implementation of said standard, as well as tips and tricks on how to best program certain aspects.

This brings us to the final category, which is about purely technical programming knowledge, which means SAS in a clinical biostatistics environment. Of course this includes external trainings about learning how to program in SAS, how to write macros and construct figures; but also very specific trainings on how to use the custom macro's, how to set-up the programming environment for a study, etc. As around 90% of my time is spent programming, it goes without saying that these topics are critical to a successful onboarding.

TIMELINE

Perhaps the best way to explain the onboarding process, is to provide a chronological overview of my first 10 weeks as a biostatistical analyst. This includes a small 'before section', followed by the first weeks, the transition to the first delivery, finishing with a short overview of what happened (and will happen) afterward.

BEFORE

Before my first working day, my manager developed an individual training plan. In short, this was done by comparing the standard function description with my CV. In addition to preparing trainings, a suitable first project is selected. In my case, this was the entire analysis of a relatively small phase I study, for a long-term client. The manager also selected my coach and assigned a desk, ensuring I was surrounded by both experienced colleagues and fellow new recruits. Finally, the secretary arranged all practical and administrative matters so the first day ran smoothly.

THE FIRST (TWO) WEEKS

The first day was mainly spent on administration, practicalities and getting to know the people: manager, secretary, coach, and, of course, direct colleagues. Most of the rest of the first two weeks was spent on external trainings to learn how to program in SAS. I spent the remaining two days training applicable SGS SOPs for my function, usually in an internal digital quality management system that ensures every employee is up to date on the latest SOP versions and trainings. A final less obvious, but very important, activity in this time period was to get to know my colleagues.

WEEK 3-4

When I was familiar with SGS procedures and SAS programming, I spent most of the next couple of weeks attending live trainings in (very) small groups. There, experienced colleagues first explained how the programming environment is set up for a study. A second (set of) trainings were an explanation about the structure of SDTM data, as well as on how to use an annotated case report form (CRF). In other words: how the input data is structured. Then, I was introduced to all the in-house macro's that are used inside (and outside) of the statistics department, as well as how they are developed and maintained. A final set of trainings was an introduction to the ADaM data structure, and more specifically the SGS ADaM implementation guide.

While attending these trainings, my first project was presented. As mentioned before, this was a relatively straightforward phase I study for a client our department worked (and works) with on a regular basis. A colleague with a lot of experience with similar studies was assigned as a coordinator to help where needed and keep an eye on the progress. The only task at this stage was to program the analysis, with the SAP already finalized.

In this phase, the majority of my time (around 80%) was spent on training sessions, with the remainder divided between finishing individual training via the digital quality management system and reading documentation about the first project (such as the protocol and SAP), as well as setting up the programming environment for this study. It is important to note that all these trainings were done face to face, in the office. This made direct interaction with the trainers and fellow recruits very easy and natural, contributing to a good training experience.

MONTH 2

After the first 4-5 weeks, I experienced a shift in both the intensity and content of the trainings.

The content shifted more from hard skills that have a direct impact on the daily work (i.e. programming) to what was referred to as 'background information' above. This included a day at company headquarters to be introduced to the general company structure, how HR works, etc. Another two days were spent on a thorough ICH-GCP training, followed by an examination. A last important training was an overview of how the biometrics division outside of the statistical department functions, making it clear how the dataflow within the division is organized, where the data comes from and what eventually happens with the analysis results.

In terms of intensity, I found myself now spending the majority of my time programming, rather than attending trainings. Of course, trainings still constituted a significant portion of my working time, but the balance was now closer to 70% programming and 30% training. Although the eventual code was fully written by me, I asked questions to the coordinator, my coach and other colleagues regularly, especially at the beginning.

FIRST DELIVERY

At the end of the second month, I delivered an internal test run of the analysis of my first study to an experienced colleague. This colleague thoroughly reviewed the result of my programming effort and gave feedback on the results. I then updated the programming where needed during weeks 9-10, after discussion with the coordinator and reviewer.

Although this step was of course primarily a quality insurance for the eventual delivery to the client, it also served as a very important learning opportunity for me. Not only in terms of purely improving programming skills or project organization, but also to learn about company procedures for delivery and review. Especially since the review process is done for most studies regardless of the experience of the people involved.

After appropriate training about delivery procedures, I sent out the result of my programming effort to the client at the end of my 10th week as a junior biostatistical analyst.

AFTER

Of course, after the first delivery, the onboarding process is far from over. Especially in terms of the 'daily work' contents, the learning virtually never stopped. First of all, the first analysis was a relatively straightforward phase I trial. Subsequent projects required more advanced SAS training on, among other things, macro programming and the creation of graphs. More complex study designs required more complicated ADaM datasets in later projects, in turn requiring more lengthy and detailed ADaM training.

Although my main task is programming, I also worked on project-related tasks such as writing an SAP, developing metadata, etc. For each of those tasks, appropriate training and documentation was available. Furthermore, there are regular refreshment and upgrade trainings about a variety of topics, ranging from ICH-GCP refreshment to the introduction of new macro's into our in-house library.

EVALUATION AND FOLLOW-UP

As mentioned in the section about the people involved, the onboarding is not a simple 'fire and forget' process, but rather a continuous interplay of training, reflection and adjustment. The follow-up happens within time windows of different sizes, ranging from almost minutes to a span of months. This allows both immediate adjustment where needed and long-term reflection and feedback.

The shortest of such windows is during or right after training sessions: it is possible (and important) to ask questions and reflect on what is being taught. Furthermore, coach and colleagues are available for immediate questions about practicalities or daily work. Similarly, study-specific questions can lead to feedback on programming code by the project coordinator and/or coach and colleagues.

Intermediate follow-up happens on a timescale measured in weeks. One such opportunity is the review of the programming described above, which happens over the course of 1-2 weeks. Additionally, an individual follow-up meeting is scheduled monthly with my manager. These regular follow-up sessions are held for all employees, not only for new recruits. During these meetings, my trainings and project progress that happened in the past month are discussed, as well as what is planned for the upcoming month (or further ahead). This allows for adjustments in the scheduling of future trainings (and projects) where necessary, but also for improvement of existing training, procedures, documents, etc.

Finally, an official evaluation meeting is conducted every 6 months. Again, these evaluations are done for all employees, although they only happen yearly after the first 1.5 years of employment. In preparation of these meetings, my manager gathers input from the study coordinator(s) and coach, in order to discuss my performance over a longer period of time. At the same time, these evaluations provide me with an opportunity to share my thoughts and feedback on the onboarding process from a more general point of view, as opposed to focusing on individual trainings or topics as is done during the monthly follow-up meetings.

CONCLUSION

The process described above was beneficial for me as a new and inexperienced employee, as well as to my managers, despite requiring a significant investment of time and effort from both sides. The reason why the process was beneficial is that I was able to contribute to the work of the department within a relatively short period of time. The approach of first focusing on daily activities and delivering trainings (more or less) at the point in time when they were most needed allowed this almost instant productivity. From my side, this 'just-in-time' approach made sure I was not overwhelmed by having to go through too many new things at once and by creating a natural order of priority. Finally, taking the time for thorough training instead of leaving me to my own devices from the start made sure that the quality of the first analysis was (relatively) high, eventually saving time and effort down the line during the review process. At the same time, applying new knowledge immediately in practice ensured the trainings did not remain dry, theoretical matter, but their use became clear shortly after the actual training.

Later on during the onboarding process, I felt it was good to learn about more background information, enabling me to situate my role in a bigger picture. This not only benefitted me personally, but also helped me in the cooperation with other departments and interaction with clients. As there was no immediate need for this knowledge at the beginning or during programming, it was good to postpone these topics till later.

Of course there are also drawbacks to this 'just-in-time' process. The first is simply practical: it is not always possible to provide a training at the needed time due to scheduling conflicts or logistical constraints. Furthermore, there is still a relatively big amount of new information to be processed in a short time by a new employee, some of which may fall through the cracks because it was not needed immediately for the first study, requiring re-training later on.

Allowing me to program an entire study as my first project had both advantages and disadvantages. On the upside, it ensured I became familiar with virtually all aspects of a project: both ADaM and output programming; different study domains; administration; study set-up, etc. However, programming an entire study by oneself presents a fairly big challenge to someone with little to no experience in SAS programming. It may be easier to start a new employee in a bigger ongoing project where (s)he can focus on, for example, only the outputs of one domain, where ADaM datasets are already programmed by a more experienced colleague. Or program an ADaM dataset when some basic datasets (e.g. ADSL) are available already. This would also allow for more direct feedback during the programming process itself by the colleague also involved in the project.

A final remark should be made about the fact that during this onboarding process, I was always at the office (during working hours). It was very beneficial for me to always be able to ask for help or input immediately and face to face, without for example having to send an e-mail that could be misinterpreted or answered only a week later. Although it is possible to work home-based for some time, I still prefer working at the office so I can discuss with other colleagues and departments directly.

That being said, my overall onboarding experience was a very positive one, allowing me to grow gradually in my role as a biostatistical analyst. The just-in-time approach contributed to my self-confidence and the quality of the analyses. Even after a short amount of time, I felt I could contribute in a meaningful way, rather than being given meaningless small tasks or, conversely, being thrown in the deep end and left to sink or swim. Whether this mirrors the feeling of my managers, colleagues and clients is not for me to say of course ...

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Contact the author at:

clinicalresearch@sgs.com

EUROPE: +32 15 27 32 45

AMERICAS: + 877 677 2667

WWW.SGS.COM/CRO

Brand and product names are trademarks of their respective companies.