Reflections on a Career in Clinical/Statistical/SAS® Programming (35+ Years in the Making)

Alan Meier, MedImmune, LLC.

ABSTRACT

Since 1982, technology, standards, submissions and environments have evolved. I've gone through mergers, job changes and re-organizations. Worked as a programmer, moved into management and returned to programming. Along the way, the 2 lessons that I learned were to be flexible and have a plan.

INTRODUCTION

The year was 1982…

- Ronald Reagan was the U.S. president; the U.K. fought the Falklands War with Argentina; the U.S. transferred the Canal Zone to Panama; the super-powers were still testing atomic weapons; the Vietnam Veterans Memorial opened in Washington, D.C.

- The Space Shuttle became "operational"; IBM released PC-DOS version 1.1; the Commodore 64 computer was released; the 1st Compact Discs (CDs) and CD Players were released; TIME's Man of the Year was a computer; AT&T agreed to divest itself of 22 Bell System companies

- The CDC reported on AIDS symptoms in 3 hemophiliacs and on a transfusion recipient who developed AIDS symptoms; the 1st permanent artificial heart was successfully implanted (Jarvic-7); the Tylenol scare occurred when 7 people ingested potassium cyanide laced capsules

- "E.T. the Extra-Terrestrial" was the highest grossing movie; "Physical" by Olivia Newton-John was the most popular single; "Thriller" by Michael Jackson was the most popular album; the 1st issue of "USA Today" was published; EPCOT Center opened at Disney World in Orlando, Florida; the Weather Channel first aired on cable television

My start in the Pharmaceutical Industry is not nearly as significant as the above events, but as there were many significant changes/events in the world, pop culture, and technology in the last 35 years, so have there been significant changes in the Pharmaceutical Industry. 21 CFR Part 11, CDISC, EDC, and e-Submissions have all had significant effects on the Clinical/Statistical/SAS Programming function.

CAREER

I started at a small subsidiary company that was combined with others into a new large pharmaceutical division and relocated. A couple years later I left for a new position, only to be merged in 6 months back with the original company. Many years later, I went to work for MedImmune only to be bought out again 2 years later by AstraZeneca.

In my career, I went from an individual contributor to nearly 20 years in management and then back to individual contributor. After the first merger, I decided that I wanted to go into management. I let my management know what I wanted and worked with them to find the opportunity to advance. Eventually I was promoted to Supervisor and then Manager. After joining MedImmune and eventually being promoted to the Associate Director level, I knew that I no longer had any desire to advance further. Later I became interested in becoming a remote employee and knew that my management probably wasn't ready to allow that, so if I wanted to stay at MedImmune, I figured that I would need to step away from management. In 2010 the opportunity to return to an individual contributor presented itself and I jumped at it even though it was several years before I actually became a remote employee.

OFFICE ENVIRONMENT

When I started, I was in a private office and from there I entered Dilbert's world of cubicles. After that I moved in and out of shared and private offices, a "glass fish-bowl" (private cubicle with floor-to-ceiling glass walls), eventually back into a cubicle, then experienced the "open-office" environment.

In 2014, I gave up all of that and became a remote employee. Today, I can look out my office window and see the Green Mountains of Vermont. My commute is less than a minute (walk up the stairs) and the only traffic I need to look out for are the puppy toys lying on the family room floor.

While work-from-home one or two days a week was an option for several years before I made the move, there was some adjustment needed to become a permanent remote employee and it may not be for everyone. Teleconferencing, video conferencing, instant messaging and e-mail all help to keep in contact with co-workers, but it's not the same as seeing them face-to-face. You don't really get the "hang around the water-cooler" time to just chat with each other. It's not as easy to stop by and thank someone for helping you out. You also need to be ready to let go of the company sponsored celebrations, parties, happy-hours, etc. If you value the benefits of remote work (no commute, more time with family, better working environment, etc.), you have to let go of the other stuff.

Most companies already offer in-house employees the option to work some days from home and with the use of Contract Research Organizations (CROs) and Functional Service Providers (FSPs), more and more work will be done outside of the corporate offices. It's cheaper for companies to do this as they don't have to build or acquire addition office space when they expand. Some open-office environments already are set up so that employees no longer have dedicated work spaces, but choose an open spot when they get to the office. They may have a small personal storage bin that they can roll with them and personal desk phones are automatically aligned as the employee moves from spot to spot. Gone (or soon to be gone) are the days of having personal items arranged on your desk and Dilbert cartoons tacked on cubical walls.

ORGANIZATION

In my 35 years, I've had at least 8 different direct supervisors and too many changes to count in the levels above me. Each change has had trickle down effects on day-to-day activities, but the closer the change to my level, the more significant the effects.

I've seen just about every permutation of Statistics, Data Management and Programming functional area combinations, with changes about every 5 years. (It's interesting to note that every time all three were combined, the department was called "Biostatistics and Data Management", with no mention of Programming.)

My early career was in "big Pharma", where making changes required months of planning and discussion. Talk of shareholder value at the working level was almost never raised. MedImmune, at least before AstraZeneca, was "Biotech start-up". Implementing change happened as needed and with minimal planning and meetings. We did what was needed to get things done as almost every day we needed to show value to our investors.

There always have been and will continue to be mergers and acquisitions in the pharmaceutical industry. In the first decade of the 21st century there were numerous "mega-mergers" where large companies merged or were acquired. This push was partially due to the desire of companies to discover and bring to market new "block-buster" drugs (greater than $1B in annual sales) after seeing the success of drugs like Viagra and Lipitor. This was also the time where biotech companies became attractive targets as new therapies from them began to reach the submission and approval stages. These activities have slowed down lately, but if pipelines start drying up, don't be surprised to see a new round begin.

TECHNOLOGY

When I started in the Pharmaceutical industry, all data and programs were loaded using a Remote Job Entry (RJE) station. Imagine a 7-3/8" by 3-1/4", stiff paper card with the upper left corner removed. Characters were recorded on the cards using a series of holes punched into the cards. Each card contained up to 80 characters (bytes) of data. One megabyte of data required a stack of 12,500 cards that's over 7.25 feet tall. If one character was mistyped, the entire card had to be re-punched.

All data and all programs had to be punched on to the cards and then loaded/executed on a mainframe computer by running the cards through a card-reader. The output came out on an impact printer on continuous form paper (usually with alternating green/white sections).

It's understandable for people who wrote programs on cards to want to pack as much as they could on a single card – especially after having to reassemble a deck that was dropped or shot out of a malfunctioning card reader. No comments, non-descript variable/data set names, no labels, no indentation and multiple statements per card were all common practices.

In 1982, disk storage was very expensive, so any data or programs that were stored (after being loaded from cards) was usually stored on magnetic tape. Floppy diskettes for data storage were available since the mid-1970's (8" version), but didn't really become popular until the advent of PCs (in 5-1/2" and 3-1/2" versions). These were subsequently replaced by hard-drives, CD/DVD technologies, flash memory and now the cloud.

Electronic cathode ray tubes (CRTs) had been around since the mid-60's and with lowered disk storage costs, we eventually transitioned to sharing a limited number of these "dumb" terminals that replaced the card readers. SAS code statements were still limited to 80 characters per line, but there was no longer the worry of dropped card decks and changes were relatively easy to make. Data entry also started to use CRTs with simple, fill-in-the-blank type data entry systems.

In the early 80's, Apple, Commodore and other personal computers were growing in popularity. Eventually IBM joined in the fray with a line of business personal computers and in 1987 came out with the PS/2 line. These were DOS based machines and there were several user interfaces available, including MicroSoft Windows. While these were "smart" machines, SAS code was still written on dumb-terminal emulation packages. These types of emulation packages are still used today, especially for batch submission of SAS programs on UNIX/LINUX environments.

With the maturing of the internet, remote desktops have come into existence. These WEB-based applications allow users to log onto corporate intranets and access internal applications from anywhere and from almost any type of machine/operating systems. These allow remote workers and CRO/FSP staff to function as if they were in-house. I expect to see these types of systems to continue advancing and including more applications in the future.

DATA COLLECTION

Paper Case Report Forms (CRFs) were the only way that data were collected in the 80's. Each page was manually created by cutting printed text and collection boxes and attaching them to sheets of paper by applying a wax coating on the backs. These packs of CRF pages were then sent to a printer for typesetting and printing. Eventually this process was replaced with form design software packages.

As technology progressed, databases were developed that eventually allowed remote data entry at investigator sites. These early Electronic Data Capture (EDC) systems were stand-alone PC-based systems that periodically downloaded data to a central database. With the maturing of the internet, WEB-based systems were developed, giving investigators and sponsors near-real-time access to clinical data.

In laboratories, Lab Information Management Systems (LIMS) were developed that were able to electronically capture lab results from auto-analyzers and transfer those to sponsors. Similarly, many types of data are now captured electronically, including Vital Signs, EKGs, Patient Questionnaires and even Drug Dispensing/Reconciliation data.

With the recent push in the healthcare industry to migrate to electronic medical records, new opportunities were created for the automatic transfer of relevant data from those systems into current eCRF systems. I'd expect see more and more remote data capture in the future.

REPORTING FORMATS

In 1982, Statisticians still used calculators for simple statistics and hand-drew figures. Reports were typed on typewriters or eventually on special purpose word processers. With PCs and Windows came MicroSoft Word and documents got passed around for review and comment. With today's authoring systems, documents are authored and shared in real time.

Regulatory submissions used to be in paper and it was not uncommon for submissions to require multiple truckloads of paper being shipped to the FDA. In 1990, the International Conference on Harmonisation (ICH), now known as the International Council for Harmonisation, was founded to address the need to quickly and efficiently bring new therapies to market. One output from this group was the Common Technical Document (CTD) in 2000 that proposed a standard format for reporting of clinical trials. While it became a required standard for the EU and Japan in 2003, it remained just a recommendation by the FDA.

In the 90's sponsors were looking at ways to submit NDAs/BLAs/ANDAs electronically to the FDA. These early eSubmissions were usually stand alone systems, including data, reports and the hardware/software necessary to access the data and reports. In 1999, the FDA issued electronic submission guidances which contained recommendations for submitting documents as PDFs and electronic data in SAS transport format. In December of 2007, these guidances were removed and the FDA encouraged sponsors to use the ICH eCTD format. As of 5-May-2017, all NDA/BLA/ANDA submissions to the FDA must be in eCTD format and are now submitted electronically, so there are no more physical deliveries to the FDA.

STANDARDS

From the above, it's believable that in 1982, standards were barely a dream. In the early 90's, I quickly became a convert of standards when I handled 80+ Healthy Volunteer (Phase I) studies that I had to create tables and listing for in a single year. I worked with our team to create standard CRFs, based on study design (e.g. single-dose dose-escalation, multi-dose dose-escalation, single-dose cross-over, etc.) from which I was able to create a standard set of table and listing templates (with standardized SAS programs) that our Medical team approved.

In 1997, the Clinical Data Interchange Standards Consortium (CDISC) was founded with a mission to develop global, platform-independent data standards. Since then a string of standards have come forth that all Clinical/Statistical/SAS Programmers should be familiar with, including SDTM, ADaM, LAB, SEND, CDASH, Terminology, and Define-XML to name a few. The FDA now requires sponsors whose studies started after Dec. 17, 2016, to submit data from these studies in the data formats supported by FDA and listed in the FDA Data Standards Catalog (including SDTM, ADaM, SEND and Define formats) for all NDA/BLA/ANDA submissions.

A group of Statistical Programmers in Europe founded the Pharmaceutical Users Software Exchange (PhUSE) group in 2004 as a forum where colleagues could get together and discuss the current and future direction of the industry. Today PhUSE is a global organization with over 8000 members and serves as an industry voice to regulatory agencies and standards organization with expertise in implementing standards. PhUSE has a number of working groups including "Optimizing the Use of Data Standards" and "Standard Analyses and Code Sharing". It holds an annual Computational Sciences Symposium (CSS) that's co-sponsored by the FDA. Deliverables have included Study Data Standardization Plans and Regulatory Reviewer's Guide templates and standardized report templates with SAS and R code to create them.

CDISC is currently working on, among other things, developing disease specific Therapeutic Area Users Guides (TAUGs). I expect that work to continue for some years. PhUSE has a lot of new and ongoing work streams. The two organizations have recently announced a change to their partnership to strengthen their interdependent process (CDISC developing the standards and PhUSE providing expertise on implementing them).

VALIDATION

In 1997 a shockwave was felt within the industry. It was the release of Title 21 CFR Part 11 – Electronic Records; Electronic Signatures – Scope and Application. Even though the European Union's Good Manufacturing Practice, Annex 11 had been out since 1993, the U.S. pharmaceutical industry mainly ignored formal validation of SAS code used to generate Tables, Figures and Listings (TFLs).

After the release, most sponsors quickly moved to various methodologies to provide a risk-based, documented approach to validation, including the now popular competing-code method where two programmers independently produce the same deliverable and compare the outputs. This nearly doubled

the workload of Clinical/Statistical/SAS Programmers overnight and sent companies scrambling to find the needed head-counts.

SAS® SOFTWARE

When I started, the company was running SAS 76 (1976) and that was quickly upgraded to 79 (1979). Pharmaceutical companies back then (and some yet today) were slow to load new versions, waiting for bugs to be found and fixed prior to loading the upgrade. In the "real world", SAS Graph® was added in 1980 as a patch to SAS 79 and macro language was added to SAS 82 (1982). With version 4 (1984), SAS switched from using the release year to a number for versions and ported SAS from IBM mainframes to UNIX and VMS. Version 5 (1986) added the interactive interface. After being rewritten in C, version 6 (1985) was the first PC DOS version and intermediate releases added Screen Control Language (SCL) and SQL. Version 7 (1998) added ODS and version 8 (1999) added Enterprise Guide®. (Aster)

Below is Table 1 which lists the major SAS releases and the approximate release dates:

Release	Date	Release	Date	Release	Date	Release	Date
72	01Jan1972	6.06	01Mar1990	8.0	01Nov1999	9.3	12Jul2011
76	01Jul1976	6.07	01Apr1991	8.1	01Jul2000	9.3m2	29Aug2012
79.5	13Apr1981	6.08	01Mar1993	8.2	01Mar2001	9.4	10Jul2013
82.4	01Jan1983	6.09	01Oct1993	9.0	01Oct2002	9.4m1	15Dec2013
4.06	01Mar1984	6.10	01Oct1994	9.1	01Dec2003	9.4m2	05Aug2014
5.03	28Jul1986	6.11	01Oct1995	9.1.3	01Aug2004	9.4m3	14Jul2015
6.01	01Jan1985	6.12	01Nov1996	9.2	01Mar2008	9.4m4	16Nov2016
6.03	01Mar1988	7.0	01Oct1998	9.2m2	01Apr2010	9.4m5	19Sep2017

Table 1. Major SAS Releases 2 (Wicklin)

In my opinion, there were 4 significant enhancements to the base SAS programming language that changed the way code was written. The order I present them in is based on release dates, not on impact. The first is SAS Graph. It has come a long way since it was first released, but it replaced hand-drawing or going to third party packages. Some say it is still too cumbersome, but that may just be due to not using it as often as other aspects of the language. Second is the macro language. This replaced SAS code that wrote SAS code to temporary external files which were then read in and executed. This greatly enhanced the ability to write generic, reusable code. Third is the REPORT procedure (another item in an intermediate release of version 6). For many years, the only way to get well formatted reports was with time consuming and inflexible FILE PRINT and PUT statements. Along with the last significant enhancement, ODS, we are now able to easily create publication ready documents.

FUNCTION

In the past 35 years, I've seen nearly every manor of organizational change (some more than once), office environments change, technology go places never thought of in 1982, data collection instruments and reporting formats change, the introduction and maturing of data, analysis and reporting standards and formalized validation, and SAS "explode" in breath, scope and functionality. As it got easier to collect data, more data was collected. As it got easier and faster to write analysis programs (from cards to shared terminals to individual PCs) and as SAS added new PROCs and functions, more analyses were done with shorter time lines. As science and statistics changed, the types of data and analysis methods changed.

What hasn't changed that much is the function of Clinical/Statistical/SAS Programming. We still write SAS code to create data sets and TFLs. Yes, there is more documentation needed today and the output formats are different, but at the core, not much has changed.

One area where sponsor company Clinical/Statistical/SAS Programmers are being asked to work differently is oversight of CROs and FSPs. For years, management has tried to deal with the workload peaks and valleys. More and more sponsors are turning to CROs and FSP agreements to level these out. While this strategy offloads programming tasks when there aren't enough internal resources, someone at the sponsor company has to oversee the work done at the CROs/FSPs. The Clinical/Statistical/SAS Programmers who are assigned as a Point of Contact (PoC) for a CRO or FSP will

find themselves spending a tremendous amount of time reviewing documents (e.g. scopes of work, specifications, TFLs, validation documentation, project timelines, performance indicators etc.). To succeed, the PoC will need to have all of the institutional and industrial knowledge needed to do the programming job, project management skills and good communication skills.

In the short-term I don't expect much to change. I expect to see new data types, especially in the area of biomarkers. There may be advances in Graphical User Interfaces (GUIs), perhaps becoming more graphical. For example, if there is mature metadata, a user might be able to drag and drop an icon of one data set on to another and the system would be able to generate the MERGE code need to join the two data sets. As computer scientists push the artificial intelligence boundaries, there may be less of a need for Clinical/Statistical/SAS Programmers, but that's still a ways off.

LESSONS LEARNED

Whether or not you like your current work environment, know that it will change. Companies merge, divest, reorganize, relocate, remodel; people come and go; science, standards and technology change. To survive, one needs to be flexible; to adapt to the changes, wait for new changes or move on. A change may be the impetuous for you to move on to something better.

If you work hard and smart and with luck, your career will go where you want it to go. But more than that, you need to know where you want your career to go and not just for the next year or two. You never know when opportunities will present themselves and in order to take advantage of them, you need to know when situations are opportunities. Two last antidotes to demonstrate this point.

The first was knowing that I was ready to get out of management since I ultimately wanted to become a remote employee. So even though it was years before I would move to Vermont, I was ready to step away from my management position when the opportunity presented itself.

Second, when I applied for my first position in the industry, I was still in school with one more year left and running out of money. I wanted a part-time job so I could complete my degree and applied for a position to mount magnetic tapes in the computer center in the evenings. The HR person called me and told me about the position I eventually took, but it was full-time. I took the job even though it would be 2 years before I could graduate. That single decision to take an opportunity set me on a path that more than 35 years later, I'm still travelling on and it's been good for me.

REFERENCES

Aster, Rick. "SAS - Versions, Environments, and Other Differences Among SAS Installations." Accessed February 6, 2018. Available at http://www.globalstatements.com/sas/differences/.

Wicklin, Rick. "How old is your version of SAS? Release dates for SAS software" The DO Loop August 2, 2013. Accessed February 6, 2018. Available at https://blogs.sas.com/content/iml/2013/08/02/how-old-is-your-version-of-sas-release-dates-for-sas-software.html.

CONTACT INFORMATION

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

Alan Meier
MedImmune, LLC.
meiera@medimmune.com