PhUSE US Connect 2018

Paper SP10

A Standardized Collaborative Approach for Statistical Programming Project Information

Harper Forbes, Hoffmann-La Roche Limited, Mississauga, Canada
Christoph Ziegler, F. Hoffmann-La Roche, Basel, Switzerland

ABSTRACT

The roles, responsibilities and key stakeholders of a statistical programmer in the pharmaceutical industry have evolved immensely over the last few decades. With globalization of resources, regulatory standard mandates, new sources of data, data sharing, advanced analytics and visualization using multiple programming languages across various software platforms, being organizationally agile is crucial to succeed in delivering data analyses. Despite the more complex landscape, statistical programmers may still be holding project information in the silo of their specific team or function, utilizing their preferred or inherited "tool" which may be subjective in location, style and structure and may be difficult to source information unless within the programming team.

In this paper we will cover some of the key considerations to building a framework for establishing a standardized, collaborative project-based one-stop shop for statistical programmers which will enable project knowledge transfer within their function and external stakeholders.

INTRODUCTION

Statistical programming project information areas are usually subjective to what the lead statistical programmer inherited, is familiar with or the preferred look of a tool. In a worst-case scenario there may not be any formalized information area, it could simply be a string of emails with information or a static document passed via email or at a certain document platform location. Even when information is formalized these repositories of details can range from restrictive online or shared network documents to internal communication media tools to websites and so on. If this information portal isn't standardized it can be difficult for new and even existing statistical programmer to access, manage, share or find information. We need to be more agile and efficient - spending time searching for information, obtaining access or updating information should be easy, intuitive and efficient. Also the "shared – sense of purpose" mentality needs to be present in order to help and to support the entire team.

BACKGROUND AND RATIONALE

In the following sections we examine some of the primary reasons for statistical programmers to consider prioritizing the creation and maintenance of a statistical programming project information area.

RESOURCING AGILITY

It is unrealistic to assume one person will lead a project from start to finish (unless perhaps an unsuccessful early phase molecule). The length of time it takes for a drug to be tested and approved can vary greatly. It might take 10 to 15 years or more to complete all 3 phases of clinical trials before the licensing stage[1]. Even in the rare circumstance the lead statistical programmer remains on a new project from start to finish of the statistical programming component the chances their team will also remain constant is very unlikely. Project peaks and dips, faster to filing efforts, early phase to late phase personnel, vendor support, changes in the resource models, all influence the personnel working on the project. Having an established approach to maintaining information allows agility and efficiency within the project for any transitions or onboarding which may occur within the team and across sites.

COMBINATION STUDIES

While a standardized approach to maintaining project information is valuable within the team directly working on the molecule it is even more beneficial between statistical programming teams when trying to source information for a molecule they need access and information on. The last 6 years has seen a dramatic increase in combination- therapy trials rising from just 13 in 2012 to 467 in 2017[2]. If teams are resourced by a primary molecule and have regulatory deliverables that may be used in combination with other therapies it should be easy for them to find and access the information that is relevant and needed.

COMPLEX LANDSCAPE

It's common for statistical programmers to work in and across many systems and platforms. Statistical programming may be run in a PC or UNIX-based environment; programming languages could vary from SAS®, R, Python™ and other languages, sometimes used together. Dynamic visualizations and data displays can be provided in applications to stakeholders. Coding can be held in a combination of stringent version controlled systems or collaborative open-

sourced environments. This can be difficult to reference and find unless held in a centralized area relevant to the project.

F.A.I.R. Data

The concept of FAIR (Findable, Accessible, Interoperable, Reusable) data principles has achieved worldwide recognition by various organizations including FORCE11, National Institutes of Health and the European Commission as a useful framework for thinking about sharing data in a way that will enable maximum use and reuse[3]. This concept is readily applicable to our own clinical trial data and should be a key mindset of statistical programmers who store, analyze and interpret data. Having findable data is one of the primary expectations of the project information repository – it provides metadata to help describe and locate the data. The accessibility of the data can be outlined with primary contact names to help make it shareable to an increasing set of stakeholders – other programmers, biostatisticians, data managers, scientists, outcome research, data curators, etc. Interoperability of the data can be outlined so users are aware of the project related agreed formats, metadata and contain links to useful areas to further project related details (i.e. Statistical Analysis Plan). Reusable data may refer to the standards implementation – if the clinical data in raw and analysis form are CDISC complaint or still maintaining a legacy format. With the increasing and understandable priority on these principles it is necessary for statistical programmers to have a framework in place for this information.

KEY CONSIDERATIONS

The background and rationale outlines a few key examples highlighting the importance of creating and maintaining a central information area to store project level information. In the authors experience, two important considerations around building an information area are to ensure the project information area is standardized and collaborative. A set structure, look and feel to allow users to navigate effectively and easily between projects while also encouraging crowdsourced, up-to-date information to ensure accuracy and efficiency.

STANDARDIZED

Similar to standardized data structures a standardized approach to maintaining project information within statistical programming may cause discomfort for some. Some lead statistical programmer may have intimate knowledge of the molecule, spending years on collecting/storing this information which they deem most convenient and efficient for them as individuals and their support team. However, knowing it's unlikely a statistical programmer will remain on the project throughout its lifecycle and the likelihood the supporting resources will change and accessibility from different stakeholders will evolve it's important to try to keep a consistent look/feel of project level information within the project and between projects. Keeping it consistent within the project (also at the study level) and between projects (different molecules) will allow programmers to navigate easily between study and project information quickly without having to familiarize themselves with a different information landscape. Instead of reactively producing a transitioning document to pass on to a statistical programmer when there's a resource change the lead statistical programmer can just simply cite the standardized information area which should contain all relevant material for an effective transition.

COLLABORATIVE

Another important consideration when building and maintaining a project information area for statistical programming is encouraging crowdsourcing the content. By implementing a central collaborative platform ("one-stop-shop", where people go first), the contributions people make create a higher level of engagement within the information area, which should promote fulfillment in helping create and maintain accurate content. Additionally, this allows the distribution of workload in creating and maintaining an area less cumbersome for one individual and shared between everyone on the team.

The authors feel if these two key themes are implemented this information area will be an ideal "one-stop-shop" for all statistical programming and is encouraged to be shared/updated by stakeholders outside statistical programming; specifically, Biostatisticians and Data Managers – who may have access to useful information which would help build the programming clinical development or "big picture" within the project information area.

ADDITIONAL CONCEPT CONSIDERATIONS

In addition to keeping the two underlying concepts – standardized and collaborative – we will now cover some of the other considerations in order to achieve successful implementation and embedding of a statistical programming information area in an organization.

ACCESSIBLE

While there could be restrictions for those outside the company or organization we recommend limited restriction within the company. This area shouldn't hold any sensitive information or results not publicized; even if there is, for example, a link to a results area containing statistical outputs these locations should already inherently be restricted from users not permitted (i.e. a document repository containing static unblinded study outputs). Information like study team members, disclosure location, study and project documentation links, etc. can all be cited. In our experience we do not recommend using the statistical programming information area for document storage as those areas should be restricted by a proper document repository.

INTUITIVE

Ensuring the statistical programming information area is intuitive is an important factor for consideration. If the tool or platform used is too difficult to build functions or too hard to find information instinctively then implementing a statistical programming information area won't be successful. Having a deep understanding of the capabilities of the primary users along with knowledge of existing tools that your colleagues are familiar with will help determine the successful adoption of the platform used for a statistical programming information area.

SUPPORTED, SECURE AND STABLE

Perhaps more of an expectation than additional concept consideration is having a system - platform, webpage(s), application or tool - that is supported, secure and stable by your organization. For support it would be ideal to have advanced knowledge of planned upgrades (this may affect the look, functionality, etc.) and a person or department to reach out to in case of technical help. The area also needs to be secure in that it should just be accessible to those in your organization or organization vendors, requiring a log-in for audit and compliance purpose. Finally ensuring it is a system where there's no threat of phasing the system out or license expirations is also important to keep in mind for long term stability.

CONTENT AND FUNCTIONALITY

Now that we've covered the key concepts and consideration for statistical programming information area we'll offer a few important content and functionality ideas to build into your central area. Functionality possibilities may also be helpful to decide on what type of platform to consider.

CONTENT

The scope of summary can be diverse – it can cover specifically the programming and analysis component of the clinical trial or – if you obtain support from your colleagues – could be end-to-end information within Biometrics. We recommend to structure the content keeping in mind there is:

- Project Specific Information: that could contain strategic information (future indications, planned filings, CDISC-implementation plan, migration to different system details, clinical development plan, etc.), deliverables that are project-driven (integrated clinical databases, periodic safety reports, etc.) and basic team information on the project level – project lead biostatisticians, data managers and statistical programmers, as well as the lead scientists, PK group, etc.

- Study Specific Information: that may cover key stakeholders on each study team relevant for a statistical programmer, links to all study documentation (SAP, Protocol, eCRF, etc.), location of programs and outputs, study details, decision logs, current tasks, etc.

The contents contained at the project and study level should be enough to serve as an ongoing transition portal should there be any resourcing changes, desire to share information in statistical programming, etc. If you find you still need to create an email or document to cover additional information this is encouraged to place in your statistical programming information area.

FUNCTIONALITY

As statistical programmers we appreciate functionality within the chosen platform which may allow us to be more efficient and agile in rendering information and allow us to explore improvement possibilities. Some of the important functionality features you may want to consider include:

- integration with other tools/websites – easy and intuitive linkage between platforms
- embed links to key documents – for all study related information (specifications, CRFs, SAP, Protocol, etc.)
- macros - ability to embed various macros for enhancement of information
- Import/export feature – importing calendars, allowing export of visuals and documents, information into static formats
- search feature – an intuitive search functionality that will search within the tool for relevant keywords
- automated emails for updates – to track updates and impact of changes; also helps ensure usage
- render documents or multimedia – documents and multimedia may be stored externally but rendered on the statistical programming information area
- mobile application – a secure mobile application promotes convenient use/maintenance
- project management – to help track key timelines and deliverables
- auditable - knowing this is a collaborative tool maintaining a history or version summary, with the ability to revert back and recover older versions, is important to mitigate any risk of losing information accidentally

When applied all of these functional capabilities help create a “one-stop shop” and should result in less time for statistical programmers finding information where traditionally they are navigating between many applications or websites.

CONCLUSION

In this paper we discussed some of the key considerations to building a framework for establishing a standardized, collaborative project-based one-stop shop for statistical programmers. This area will reduce time required for project knowledge transfer within their function and to external stakeholders, increase transparency around information and help enable everyone using the area to understand the “big picture” while reducing time spent storing this information in various areas by establishing a single area for reference.

While this paper commonly referred to statistical programming information area there is the ability for this to be used across all Biometrics to serve as a Biometric “one-stop-shop”. Statistical programmers are usually the function in-between Clinical Data Management and Biostatistics and are thus familiar with both the collection and analysis needs of the study. Much of the relevant information contained in the information area is also critical for these stakeholders. We encourage those willing to adopt this information area to leverage the content with other stakeholders to increase accuracy, maximize crowdsourced information and reduce parallel information capturing or overlap in Biometrics.

Primary difficulties establishing this framework may be around the tools available and change mindset. While the former may be subjective to the organization, showing the value in establishing a statistical programming information area, along with proper change management, including leadership endorsement, can be effective to allow successful adoption.

REFERENCES

1. Cancer Research UK. 2015. How long a new drug takes to go through clinical trials. [ONLINE] Available at: http://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/how-clinical-trials-are-planned-and-organised/how-long-it-takes-for-a-new-drug-to-go-through-clinical-trials
2. Jun Tang and Vanessa Lucey, Anna-Maria Kellen Clinical Accelerator program, Cancer Research Institute.
3. Australian National Data Service. 2018. The FAIR data principles. [ONLINE] Available at: https://www.ands.org.au/working-with-data/fairdata

CONTACT INFORMATION

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

Harper Forbes
Hoffman-La Roche Limited
7070 Mississauga Road
Mississauga, ON L5N 5M8 Canada
harper.forbes@roche.com

Christoph Ziegler
F. Hoffmann-La Roche
Hochstrasse 16
4053 Basel, Switzerland
christoph.ziegler@roche.com