Paper 116-2023

Another Glance at Good Programming Practice from the Perspective of FDA Reviewers

Hongbo Li, Amylyx Pharmaceuticals, Inc.

ABSTRACT

Statistical programs/codes are part of submission packages to regulatory agencies for new drug/biological product applications in the pharmaceutical industry. In the PharmaSUG 2023 conference at San Francisco, one of the FDA panel presentations did raise some critical issues about the submitted SAS programs based on FDA reviewer's perspective, for example, lack of sufficient comments, lack of necessary information in the program header, not all macros called in the program have been submitted, all analysis performed in one program, lack of usefulness. Therefore, it is essential that the statistical programmers in this industry follow the guidelines of Good Programming Practice (GPP).

PhUSE GPP Steering Board is a voluntary industry group with representatives from a diverse array of health and life sciences organizations, who had developed a guidance on GPP in March 2014 which could be found on phUSE Wiki. It highlights the main standards that should be followed and what should be avoided. However, it seems that this GPP guidance is not well spread out or raised attention to the entire industry so that the FDA reviewers have found the above common issues which could be easily avoided if the programmers follow the principles and coding conventions of GPP.

This paper will summarize the essentials of GPP combining the above issues aroused by FDA reviewers, illustrate the fundamental principles and coding conventions of GPP, clarify What to Do/Don't in the program and explore the appropriate layout and header format as well as showing the examples of good and bad codes. It will be helpful to all statistical programmers, especially the beginner programmers, and also the managers who would like to embed GPP into SOP or Work Practice documentation.

INTRODUCTION

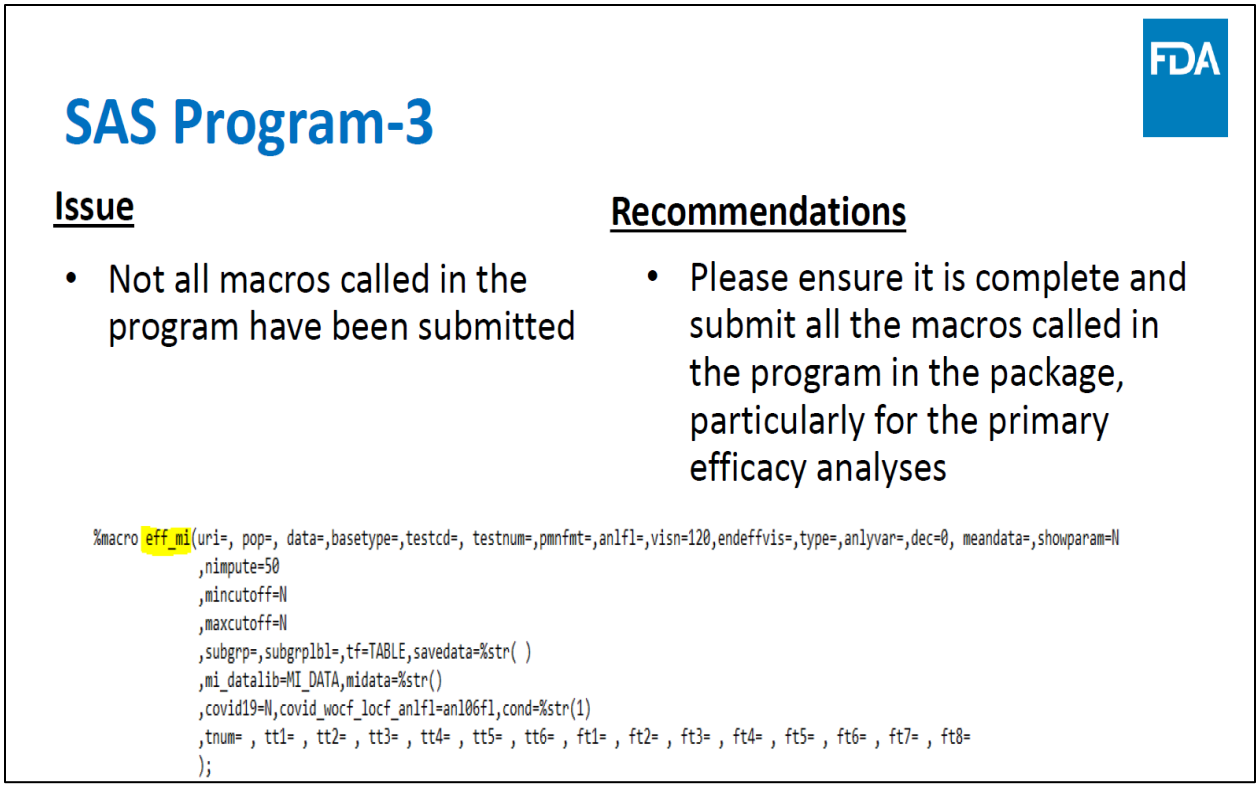
Statistical programming plays an important role in the clinical trials of drug development. Programing codes are required to be included in the submission packages for regulatory agencies' review for pharmaceutical or biotech companies' NDA/BLA submission.

In the PharmaSUG 2023 conference at San Francisco, one of the FDA panel presentations raised some critical issues about the submitted SAS programs based on FDA reviewer's perspective. As an audience, I immediately connected this in my mind with the topic of Good Programming Practice (GPP), which PhUSE has developed a very useful guidance document in 2014, available on the PhUSE website.

In this paper, I will use PhUSE GPP guidance framework and combine with FDA's presentation at PharmaSUG 2023 conference to overview some key elements of Good Programming Practice in our pharmaceutical industry.

MACRO

If macros are used in the program, make sure to include all the macros called in the submitted package. Below (Display 1) is one of the issues raised in the FDA panel presentation.



Display 1. SAS Program Issue 3 in FDA panel presentation

For the saved macro program, the file header should document the invocation parameters and any global macro variables that are created or modified. And also, the use of keyword parameters is encouraged (not required) over positional parameters. This can make it explicitly clear what values invocation parameters are being assigned and avoid mixing up with position ones.

PROGRAM HEADER

No matter which company or organization you work for, a standard program header should always be used in the program. The elements in the header should have (but not limited to): Project name, program name and location, author, date, short description of the program purpose, output name and location, list of macros used, revision history (version number, author, date and purpose of change). Other elements like input datasets or files, platform and operating system, as well as software version are highly recommended to be included in the header too.

If there is any dependency involved, ensure to include brief description in the program header to be explicit and informative to people who might read and use it as FDA reviewer recommended below in Display 2.

SAS Program-2

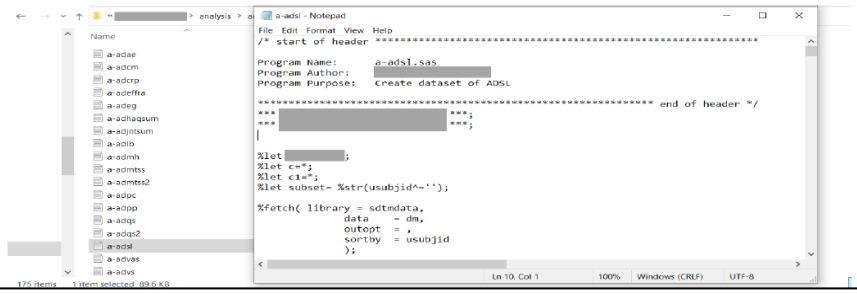
Issue

Lack of necessary information in the program header

- Missing SAS macro dependency

Recommendations

- Provide comprehensive information in the program header, such as the datasets used and the names of all macros called in the program
- Ensure the program's structure and dependencies are easy to understand



13

Display 2. SAS Program Issue 2 in FDA panel presentation

COMMENTS

Writing comments within programs is very important to help anyone reviewing, modifying, or using the program to easily understand the codes. Lack of sufficient comments is also one of the issues pointed out by FDA reviewers. See Display 3 below.

SAS Program-1

Issue

Lack of sufficient comments

- SAS programs/R scripts do not follow good programming practices

Recommendations

Having well-commented code to explain what it is doing

- Describing the purpose of the code in the header of the program
- Separating the code into chunks with a one-line description at the beginning of each chunk
- Adding comments is a simple action sponsors could take to make our work easier

Display 3. SAS Program Issue 1 in FDA panel presentation

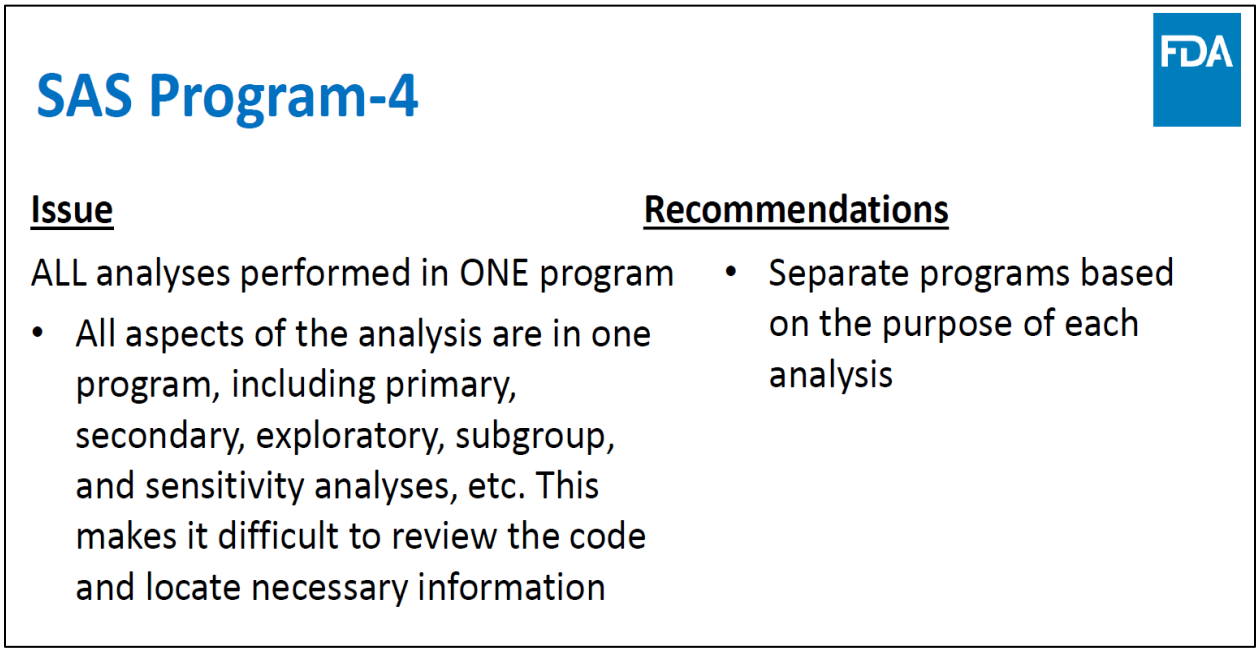
All major data steps or proc steps should be commented. Comments should be with sufficient details to allow the reader to understand what the code or procedure is doing by just reading the comments. Any complex logic or PROC SQL sub-queries should be described comprehensively with some rationale explanations. Comments can also include links to external documentation (data/output specification, design documents etc.).

NAMING CONVENTIONS

Usually, each company or organization has some standard naming conventions on program names and output names. For example, table program and output can begin with "T", followed by table number, data, analysis type and population. They could share common elements, providing meaning and ease in finding files. It makes it possible to identify groups of related programs, such as adverse event tables and listings.

Datasets and variable names should describe as best as possible their content to enhance readability and understanding, and also follow CDISC standard requirements.

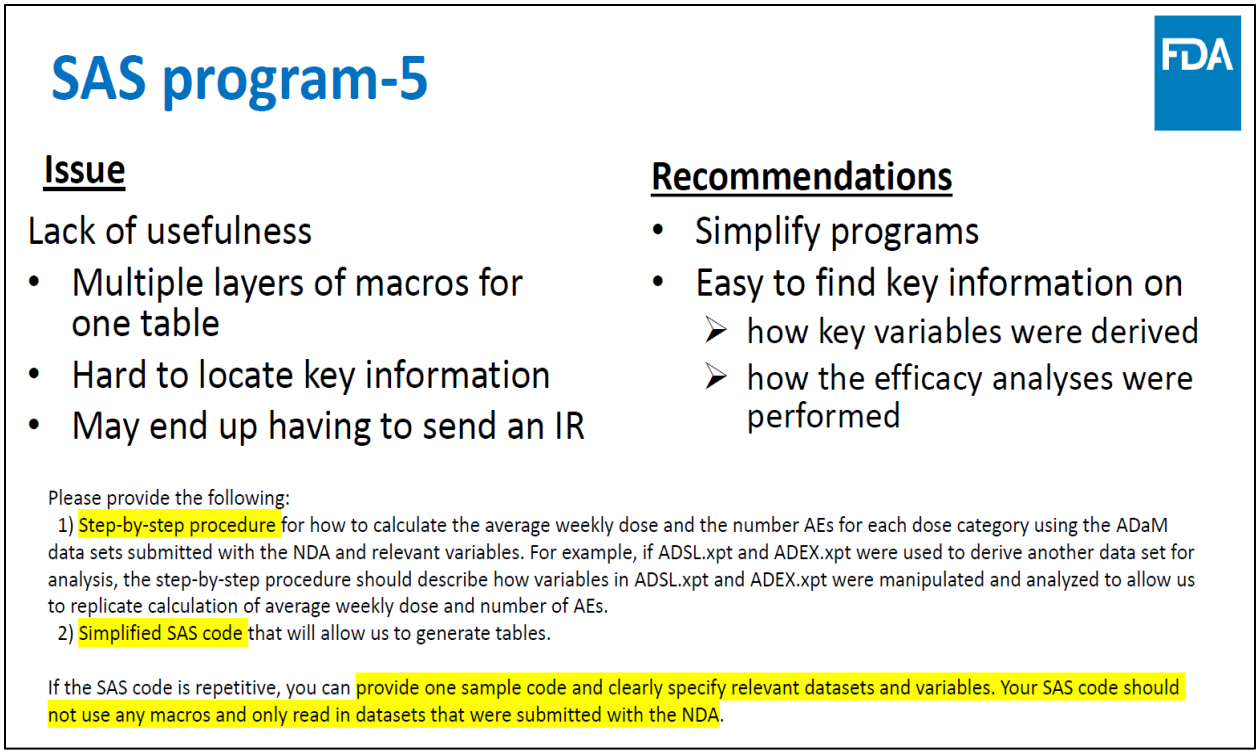
It is better that one program should not perform too many analyses together. See the below recommendation by FDA reviewers in Display 4.



Display 4. SAS Program Issue 4 in FDA panel presentation

CODING CONVENTIONS

Before talking about coding conventions, let's look at one issue that FDA reviewers indicated in Display 5 below.



Display 5. SAS Program Issue 5 in FDA panel presentation

The readability and quality of the code are important as well as easy to understand/implement. The GPP guidance of PhUSE has a full list of coding conventions, which I will not state one by one in this paper. You can find them all in the PhUSE Wiki (the link will be provided in the references). These involve a variety of coding guidelines: indentation, use of upper case/lower case, blank lines for separation, not overwrite existing datasets/temporary datasets, RUN/QUIT statements required for all data steps and procedures, etc. All these conventions are beneficial between programmers, with regulatory agencies, and with external vendors. Following these conventions can make the codes easy to read, maintain and update. Therefore, the coding conventions are vital for programmers to learn and follow.

HARDCODING AND DEFENSIVE PROGRAMMING

Programmers should be aware of the fact that they should not change data values within their programs in order to comply with the FDA's 21CRF Part 11 Rule which requires changes to source data should be done in data entry or capture systems and a clear audit trail of all data changes is needed. Writing programs to be reusable and in general situations no matter what study treatment or design is.

If hardcoding is unavoidable, clear documentation should be made: place comments in the program header and around the code, use PUT statement to the log to show it and document it in the Note to File.

Defensive programming is an approach to anticipate future changes of the data that might influence the coding algorithm during the programming process. Allow program to handle missing values, check for questionable values, process for identifying and reporting potential data issues at a later point.

CONCLUSION

In the pharmaceutical industry, we are more and more standardized and professionalized. We have various industry standards, like GCP, GMP, GLP. Of course, for statistical programming, we now have GPP also. Understanding those guidelines and implementing them in our day-to-day programming work is the key to produce the well-structured, accurate and high-quality programs, which benefits not only your co-workers, external vendors/clients, but also the regulatory agency reviewers who are decision makers on your entire submission packages. Therefore, we programmers should pay full attention and follow the guidance of good programming practice.

REFERENCES

1. PhUSE Good Programming Practice Steering Board. "PhUSE Guidance on Good Programming Practice by the Steering Board for Good Programming Practice in Health and Life Sciences, Version 1.1 March 2014." https://phuse.s3.eu-central-1.amazonaws.com/Advance/Templates/Guidance+on+Good+Programming+Practice.docx
2. Liping Sun, Statistician, CDER OTS OB DAI. "Experiences of CDER Statistical Analysts with NDA/BLA Reviews: Some Helpful Tips for Sponsors". PharmaSUG 2023 Conference Proceedings: FDA-G004. https://www.pharmasug.org/proceedings/2023/FDA/PharmaSUG-2023-FDA-G004.pdf
3. US Food and Drug Administration. "Guidance for Industry Part 11, Electronic Records; Electronic Signatures — Scope and Application." http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm125125.pdf

CONTACT INFORMATION

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

Hongbo Li
Amylyx Pharmaceuticals, Inc.
Elaine_Li@amylyx.com
+1-317-219-8433