

-
- **Aadesh Shah** is an experienced Pharmaceutical/CRO high-performance leader with a demonstrated history of excelling in the leadership, transformation and elevation of large global statistical/clinical programming organizations. He has over 13 plus years of experience working with different CRO like Octagon, ICON and Covance, he is currently working with GSK within the Biostatistics department based in Upper Providence PA. This is Aadesh's 2nd paper he has authored and presenting.
 - **Susruth Ganoothula** is an experienced statistical programmer in the pharmaceutical research and development industry who is currently working with GSK as a Principal Programmer based out of Upper Providence PA. He has over Six years of experience with a wide variety of experience working across both Pharma and CRO based working styles with companies like Pharmacyclics(Abbvie) ,Biomarin ,Receptos and Covance on different Therapeutic areas along with a Masters degree in Applied Computing. This would be his second paper at PHUSE adding to the paper which was presented last year at Baltimore.

Documentation, Applicability and Automation How Best Practices Can Help You Achieve Best Results

Aadesh Shah
Susruth Ganoothula

Disclaimer



- This disclaimer informs that the **views, thoughts, and opinions expressed by us belong solely to the authors, and not necessarily to the author's employer, organization, committee or other groups or individuals.**



Agenda



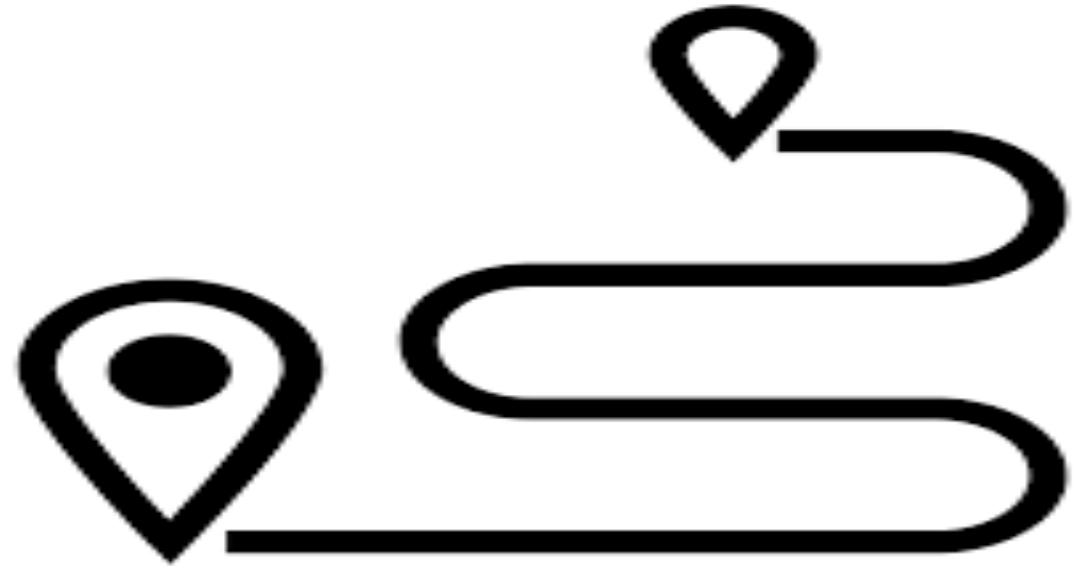
-
- Documentation
 - Traceability
 - Validation
 - Quality Gate
 - Automation
 - Human Supervision
 - Conclusion

- **What to document**
- **How much to document**

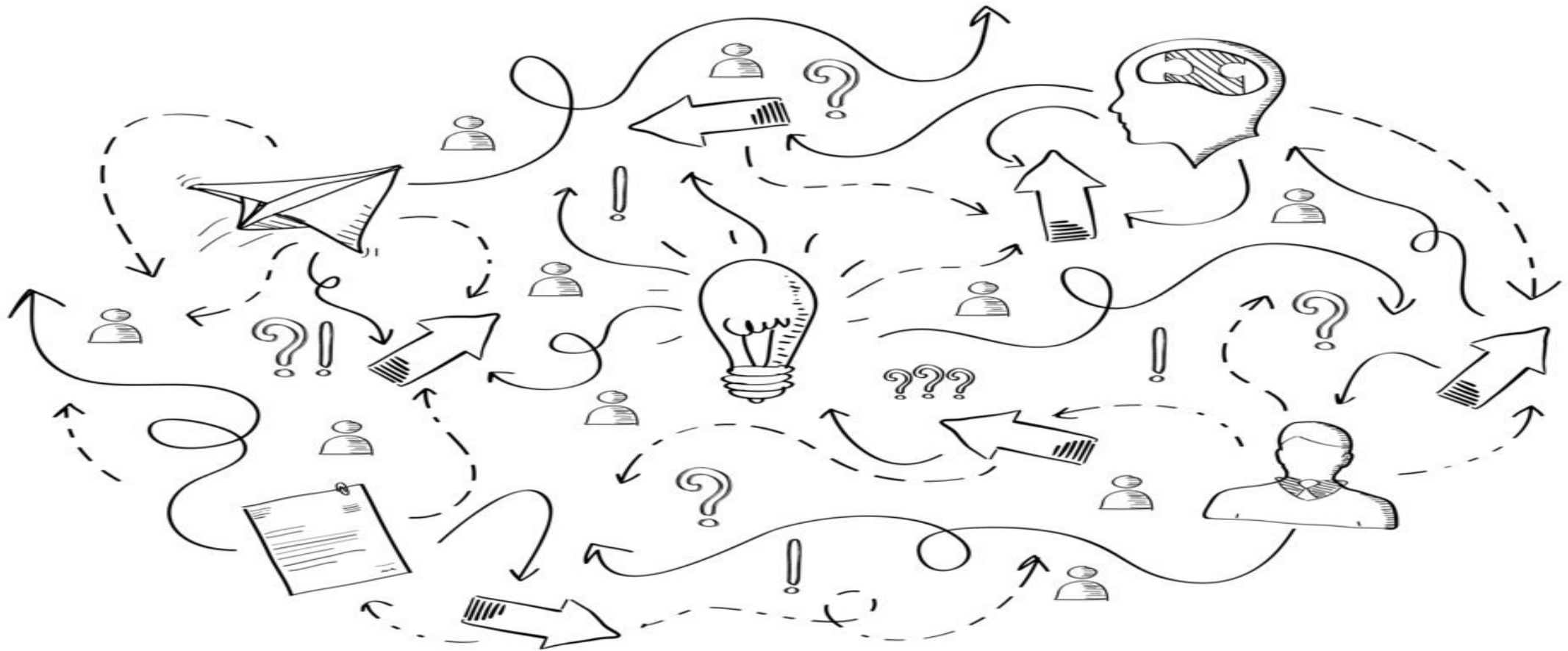
Not Documented, Not Done



- **Metadata Traceability**
- **Data Point Traceability**
- **Report Traceability**



Validation



Quality Gate



Validation Checks	Date/Time Stamps
QC Spreadsheet Check	
Confirm appropriate raw/external data have been used.	ECG – 08/29/2019
Record type and date(s) of raw/external data.	RAW – 08/29/2019
	LAB – 08/29/2019
SDTM datasets/SDTM define.xml have been updated.	Prod – 08/30/2019 – 2:48 – 2:56 PM
Confirm date/time stamps occur after date/time stamps of raw data.	QC - 08/30/2019 – 3:15 – 3:18 PM
P21 Report run/commented	
Analysis datasets/ADaM define.xml have been updated.	Prod: 7Sep2019 1:01 – 3:06 PM
Confirm date/time stamps occur after date/time stamps of SDTM	QC: 7Sep2019 4:08 - 4:10 PM
P21 Report run /commented	
Table outputs have been updated.	Prod: 8Sep2019: 2:41 -3:04 PM
Confirm date/time stamps occur after date/time stamps of analysis datasets.	QC: 8Sep2019: 3:30 - 3:32 PM
Listing outputs have been updated.	Prod: 8Sep2019: 3:04-3:07 PM
Confirm date/time stamps occur after date/time stamps of analysis datasets.	QC: 8Sep2019: 3:37-4:34 PM
Figure outputs have been updated.	Prod: 8Sep2019: 2:41-4:10 PM
Confirm date/time stamps occur after date/time stamps of analysis datasets.	QC: 8Sep2019: 4:32 PM
Logs are clear of warnings and/or notes have been explained.	
Confirm logs clear or all warnings/notes have been adequately explained.	
Archive Performed	15-Sep-19
Comments:	

Automation



Automation is by which a process or procedure is performed with minimal human assistance with the help of technology.

- **Improves Efficiency**
- **Improves Quality**
- **Optimal utilization of Resources**



Automated scanning of PROC COMPARE outputs:

- Reads all .LST files in path provided
- Looks for discrepancies, if any
- Checks if the number of variables or observations read from datasets are equal or not.
- Looks for warnings related to duplicated observations.
- Looks for variable attribute differences

Automated scanning of Log files:

- Reads all .log files in path provided
- Looks for Errors Warnings that are generated in the log
- Checks if the number of variables or observations read from datasets are equal or not.
- Looks for warnings related to duplicated observations.
- Looks for variable attribute differences
- Searches for some special notes based on keywords.

Automated scanning of Log files:

```
if indexw(upcase(_infile_), 'WARNING:') > 0 or
index(upcase(_infile_), 'ERROR') > 0 or
indexw(upcase(_infile_), 'MERGE') > 0 or
index(upcase(_infile_), 'UNINITIALIZED') > 0 or
index(upcase(_infile_), 'CATALOG IS OPENED FOR READ ONLY') > 0 or
index(upcase(_infile_), 'FORMAT WAS TOO SMALL') > 0 or
index(upcase(_infile_), 'VALUES HAVE BEEN CONVERTED TO') > 0 or
index(upcase(_infile_), 'MISSING VALUES WERE') > 0 or
index(upcase(_infile_), 'IS ALREADY USED OR NOT A VALID SAS NAME') > 0 or
index(upcase(_infile_), 'IS ALREADY ON THE LIBRARY') > 0 or
index(upcase(_infile_), 'DIVISION BY ZERO') > 0 or
index(upcase(_infile_), 'MATHEMATICAL OPERATIONS COULD NOT') > 0 or
indexw(upcase(_infile_), 'INVALID') > 0 or
indexw(upcase(_infile_), 'OVERWRITTEN') > 0 or
index(upcase(_infile_), 'AT LEAST ONE W.D FORMAT') > 0 or
indexw(upcase(_infile_), 'TRUNCATED') > 0 or
index(upcase(_infile_), 'THE QUERY REQUIRES REMERGING SUMMARY STATISTICS BACK') > 0 or
index(upcase(_infile_), 'OUTSIDE THE AXIS') > 0 or
index(upcase(_infile_), 'OBSERVATIONS WERE DELETED') > 0 or
index(upcase(_infile_), 'REPEATS OF BY VALUES') >
```

Additional Checks:

- Though logs are clean the final results might be corrupted.
- Check the logs to see if the number of observations and variables in key areas of the code to make sure the logic is working smoothly.
- Be open to defensive programming to make sure code can handle issues that may arise later.
- Use SAS procedures like `freq` or `proc sql` to make sure the counts are accurate especially while working on lab datasets.

Conclusion



-
- Maintaining quality while doing independent programming on dynamic data with lots of variables is always tricky and it is imperative that we identify and come up with crafty solutions where we are assured about the quality of the work done as a programming team.
 - We are confident that this presentation is a step in the right direction and help the programming community to implement them and achieve better quality outputs.
 - Having said that to achieve ideal programming practices we need to constantly evolve and adapt to the new challenges in the industry and come up with effective methods and revisit the current practices and analyze their efficiency and the organizations must make it a priority to train the personnel and make adherence to qualitative programming a top metric.

Questions



- Aadesh Shah
- aadesh.x.shah@gsk.com
- Susruth Ganoothula
- susruth.x.ganoothula@gsk.com

