What are the Quality Standards for Exploratory Analyses?

Kostas Sechidis, Mark Baillie, Björn Bornkamp

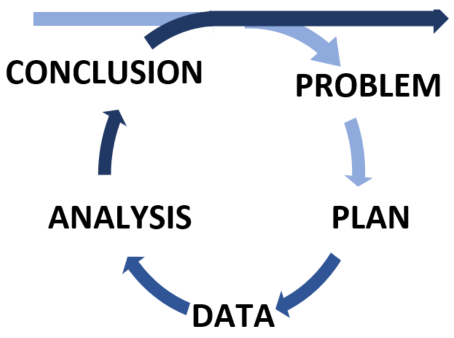
What are the quality standards for exploratory analyses from a regulatory perspective?

Statistical thinking is the foundation for empirical research.[1]
Structured frameworks ensure **clarity**, **reliability**, and **replicability** in data analysis and decision-making.

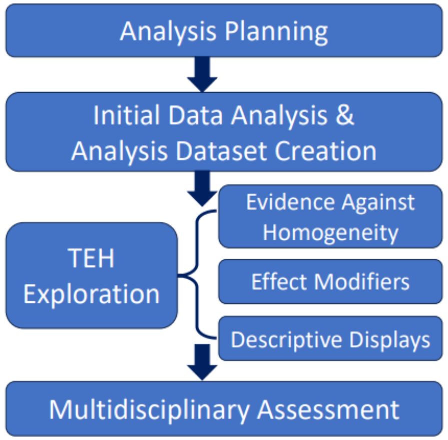
PPDAC (Problem, Plan, Data, Analysis, Conclusion) [2,3] is an effective framework that embodies statistical thinking and provides a structured approach to problem-solving by guiding researchers through each step:

- P: Pre-specification of questions and translation to target estimand
- P: Aligning the question with the data and analysis strategy[4]
- D: Initial data analysis, understanding the context and pitfalls
- A: Reproducible and accurate implementation
- C: Clear and transparent reporting

PPDAC



WATCH



The PPDAC cycle is versatile and applies to both exploratory and confirmatory analyses.

- **Estimand framework** can be seen as an example of the PPDAC cycle
- **WATCH** exemplifies PPDAC for exploratory assessment of **treatment effect heterogeneity**[5]

Why are we asking the panel this question?

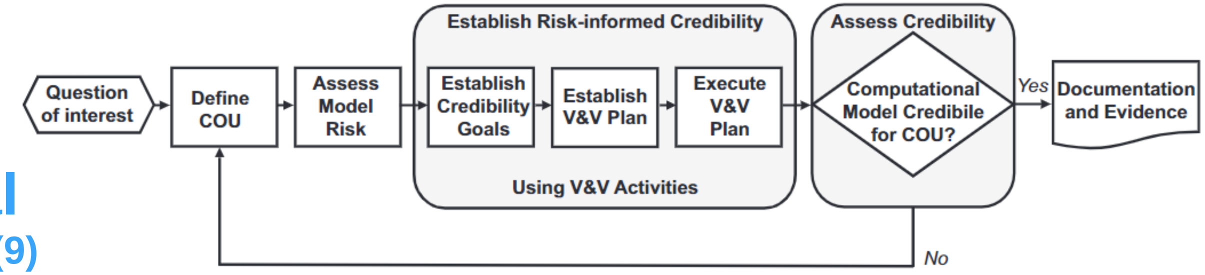
Exploratory analyses address crucial questions **in drug development**, influencing decisions on drug approval and labeling, for example
- identification of prognostic (super)covariates for adjusting (covariate adjustment)
- assessing treatment effect heterogeneity[(6)],
- safety assessments

Exploratory analyses, whether pre-specified or not, provide **essential flexibility** for exploration and learning. However, this flexibility comes at the price of requiring increased **self-discipline and rigor** to ensure high-quality outcomes.[(7)]

While there are well-established standards for primary analysis, **guidelines for exploratory analysis are less clear**, though important decisions are often based on them.

We see in other areas regulatory interest:

- In pharmacometrics the **MIDD**[(8)]
- In ML/AI the **Good ML Practice for Medical Device Development: Guiding Principles**[(9)]



With this context, we would like to open a discussion on

... **what are the Quality Standards for Exploratory Analyses?**

References

(1) Tong, C. (2019). Statistical Inference Enables Bad Science; Statistical Thinking Enables Good Science. The American Statistician.
This paper emphasizes the importance of statistical thinking in understanding and addressing complex real-world problems.

(2) Wild, C.J. and Pfannkuch, M. (1999), Statistical Thinking in Empirical Enquiry. International Statistical Review
This paper emphasizes the importance of the PPDAC cycle (Problem, Plan, Data, Analysis, Conclusion) in fostering statistical thinking and improving statistical education.

(3) Spiegelhalter, D. (2019). The art of statistics: Learning from data. Penguin UK.
Throughout the book, Sir David Spiegelhalter illustrates the PPDAC cycle using various diverse real-world question

(4) Mallows, C. (1998). The Zeroth Problem. The American Statistician
This paper introduces the concept of the "zeroth problem," which involves determining how the data relate to the problem at hand and identifying what other data might be relevant.

(5) Sechidis, K., Sun, S., Chen, Y., Lu, J., Zang, C., Baillie, M., Ohlssen, D., Vandemeulebroecke, M., Hemmings, R., Ruberg, S. and Bornkamp, B. (2024). WATCH: A Workflow to Assess Treatment Effect Heterogeneity in Drug Development for Clinical Trial Sponsors. arXiv preprint arXiv:2405.00859.
The paper introduces the WATCH framework, which utilizes the PPDAC cycl to systematically assess treatment effect heterogeneity in clinical drug development.

(6) Bornkamp B, Zaoli S, Azzarito M, et al. (2024), Predicting subgroup treatment effects for a new study: Motivations, results and learnings from running a data challenge in a pharmaceutical corporation. Pharmaceutical Statistics.
This paper describes a Novartis internal data challenge focused on subgroup identification for clinical trials

(7) Baillie, M., Moloney, C., Mueller, C. P., Dorn, J., Branson, J., & Ohlssen, D. (2022). Good Data Science Practice: Moving Toward a Code of Practice for Drug Development. Statistics in Biopharmaceutical Research.
This paper discusses the growing interest in data science within the pharmaceutical industry and proposes a framework for good data science practice in the context of drug development

(8) Kuemmel, C., Yang, Y., Zhang, X., Florian, J., Zhu, H., Tegenge, M., Huang, S.-M., Wang, Y., Morrison, T. and Zineh, I. (2020), Consideration of a Credibility Assessment Framework in Model-Informed Drug Development: Potential Application to Physiologically-Based Pharmacokinetic Modeling and Simulation. CPT Pharmacometrics Syst. Pharmacol.
This white paper discusses the application of a credibility assessment framework to Model-Informed Drug Development (MIDD), highlighting its potential to standardize regulatory evaluations and improve the reliability of physiologically-based pharmacokinetic models.

(9) US Food and Drug Administration. (2021). Good machine learning practice for medical device development: guiding principles. The US Food and Drug Administration: Washington, DC, USA.
The work outlines 10 guiding principles for GMLP to ensure the development of safe, effective, and high-quality medical devices that utilize AI/ML.