Implementation of Bayesian Quantitative Decision Making to early drug development

Gautier Paux[1], Nicolas Bonnet[2]
[1] Sanofi R&D, Vitry-sur-Seine, France, [2] Sanofi R&D, Montpellier, France

Introduction

Context

- Go/No-go decisions are critical steps where significant decisions can have a major impact on the rest of the development and the company portfolio
- ASA's statement on p-values *"Practices that reduce data analysis to **mechanical "bright-line" rules** (such as "$p < 0.05$") for justifying scientific claims or conclusions can lead to erroneous beliefs and **poor decision making**"*
- Proof-of-Concept (PoC) trial design and decision-making process should be aligned
- Frewer *et al.* (2016) proposed an alternative approach for decision making based on Lalonde *et al.* (2007) criteria
- Fisch *et al.* (2015) also provides an alternative to statistical significance

3-Outcome Decision Making framework

- 3ODM framework based on **Go/Consider/No-go** decisions provides a quantitative framework more in line with the traditional approach for decision making
- It overcomes the discouraged bright line decision approach based on a single threshold
- Decisions integrate the concept of **clinically** and **commercially** attractive effect sizes
- In this poster, we discuss key features of the implementation of 3ODM for PoC studies with revised decision rules compared to Lalonde *et al.* (2007)

Key implementation features of 3ODM for PoC studies

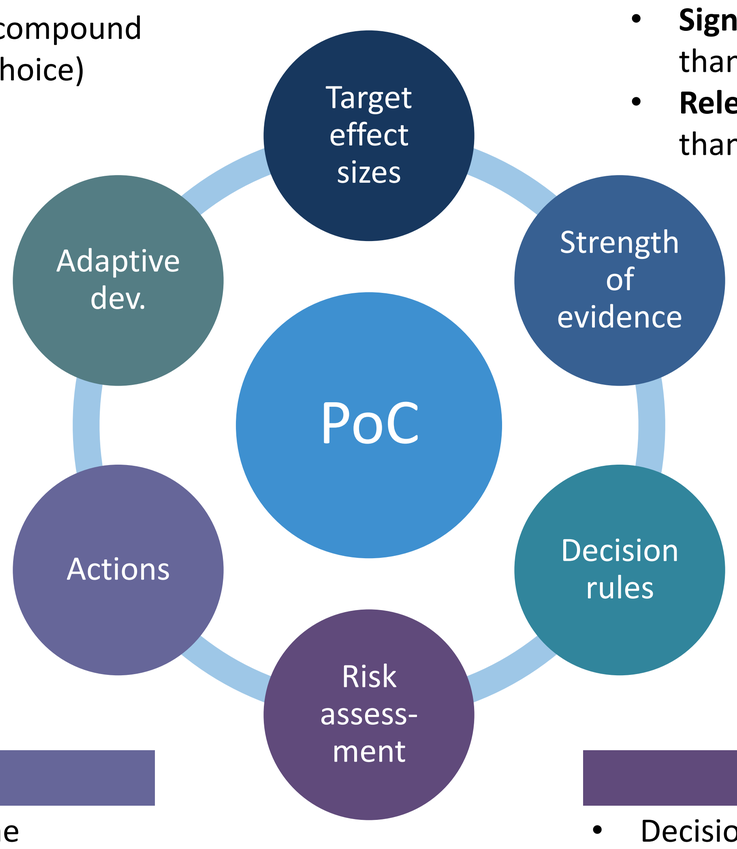
1. Target effect sizes

Lower Reference Value (θ_{LRV})

Minimally **clinically significant** effect size (i.e.., below which the experimental drug is not considered for further development)

Target Value (θ_{TV})

Minimally **commercially attractive** effect size (i.e., desired effect to potentially establish the compound as the treatment of choice)



2. Strength of evidence

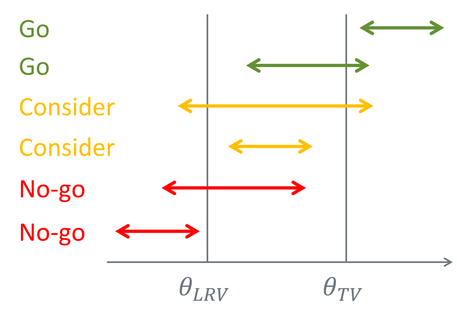
- Strength of evidence relative to the target effect sizes are assessed through a dual criterion quantifying the **significance** and **relevance** of the treatment effect, with regards to the target effect sizes:
  - **Significance criterion:** confidence that the true effect size is larger than LRV, i.e.,
  - **Relevance criterion:** confidence that the true effect size is larger than TV, i.e.,

3. Decision rules

- Go/Consider/No-go decisions are based on the comparison of the **Bayesian posterior probabilities** $P(\theta > \theta_{LRV}|X)$ and $P(\theta > \theta_{TV}|X)$ (where X are the data) to relevant thresholds τ_{LRV} and τ_{TV} respectively

$\leftarrow (1-\tau_{LRV})$-percentile $PCT_{1-\tau_{LRV}}$
Must be above θ_{LRV} to meet the significance criterion

$\rightarrow (1-\tau_{TV})$-percentile $PCT_{1-\tau_{TV}}$
Must be above θ_{TV} to meet the relevance criterion



4. Risk assessment

- Decision rules are assessed through **operating characteristics** under several true effect size assumptions
- Depending on the project priority, ranges of risk are considered acceptable by the project team
- **Quick kill**: focus on P(No-go|LRV) and P(Go|LRV) for lowest priority projects
- **Quick win**: focus on P(Go|TV) and P(No-go|TV) for highest priority projects

5. Actions

- **Clear actions** should be identified for each possible outcome
- In particular, when the outcome falls in the "Consider" zone
  - Specify key secondary endpoint(s) and or/ subpopulation(s)
  - Specify alternative development plan
- Risks are also put into the perspective of defined actions (cost, FTE)
  - Higher risks are acceptable for less costly actions

6. Adaptive development

- **Informal interim looks** can be planned to adapt the development strategy in case of overwhelming or futility results
- For example, project activities (CMC, new studies,…) can be started earlier (postponed resp.) in case of promising (unpromising resp.) results
- Interim decision rules can be based on the **Bayesian predictive probability** of a Go decision

Other considerations

Sample size

- Sample size justification does not rely on statistical significance framework anymore
- Sample size should ensure **good operating characteristics** and **good estimate precision** (e.g. width of 90% credible interval)
- **Utility function** can be used to compare several sample sizes, e.g., maximizing the weighted probability of correct decision under LRV, TV and at the mid-point between TV and LRV

Leverage prior information

- Objective is to maximize the use of existing knowledge
- In Bayesian setting, **informative prior** could be used to decrease the uncertainty and/or decrease the sample size
- Unbalanced design could be envisaged when informative prior is specified for the control arm
- Prior specification has an impact on the risk assessment and needs to be taken into account

Conclusion

Key messages

- Change in decision-making paradigm
- Quantitative tools are used to assess the **significance and relevance of effect sizes** with regards to clinically and commercially targets
- "Consider" outcome is used when results are not compelling enough; **consistent with the traditional decision-making** approach which is often not binary
- Education, communication and standardization are key to ensure a good implementation of this new framework

Software development

- Internal development of an interactive web-based application using RShiny
- Interactive assessment of decision rules and sample size by the project team
- Flexibility in terms of trial endpoint, decision rules and paradigm (Bayesian or frequentist)
- Standardized methodology, tool and outputs across project to ease the review by senior leaders

References

Lalonde *et al.* (2007). Model-based drug development.
Frewer *et al.* (2016). Decision-making in early clinical drug development
Fisch *et al.* (2015) . Bayesian Design of Proof-of-Concept Trials

SANOFI