ISPOR-EU, 2023

Acceptance Code: MSR93

# Hype, not Hope – Using Bayesian Methods in Drug Development

Wilmar Igl[1] & John Constant[2]

[1] ICON PLC, Sweden; [2] ICON PLC, Canada

## Introduction

- Recent publications advocated the use of Bayesian statistics in drug development (Ruberg et al., 2023):

  - conceptual arguments
  - recommendations for Bayesian models
  - examples of empirical studies
- Position papers used similar arguments in the past but have not impacted the adoption of Bayesian methods as much as desired.
- Issues in the used arguments

## Key Issues

- Overstatements: Bayesian statistics is not fundamentally distinct with enormous consequences (cf Gill, 2011)
- Miscommunication to target audience: Regulators' priority is to be conservative and consistent and not innovative (cf Campbell, 2020)
- Inappropriate Bayesian methods: Bayesian methods may show massive inflation of risk for false positive conclusion (cf Berry et al., 2013; Chu & Yuan, 2018)
- Poor empirical examples: questionable assumptions, no added value (cf Senn, 2022)

## Conceptual Arguments

**Fundamental differences and enormous consequences**

- Differences are overstated, because Bayesian and frequentist approaches will give the same results:
  - if a non-informative prior distribution is used
  - if more data are collected.
- Argument will scare away regulators, because the priority for regulators is more on being conservative and consistent and less on being innovative.

**Use of existing empirical data**

- Against principle of "Design trumps Analysis!"
- Prospective Randomized controlled trials (RCTs) are still the reference standard
- High complexity will raise concerns by regulators by adding hard-to-understand Bayesian methods to hard-to-understand data

## Conceptual Arguments (cont.)

**Use of existing knowledge**

- Subjective belief in the validity of a prior distribution needs to be agreed upon with regulators
- Bayesian statistics does not solve the problem of subjectivity but moves it to other locations in the decision process

**Costs and complexity**

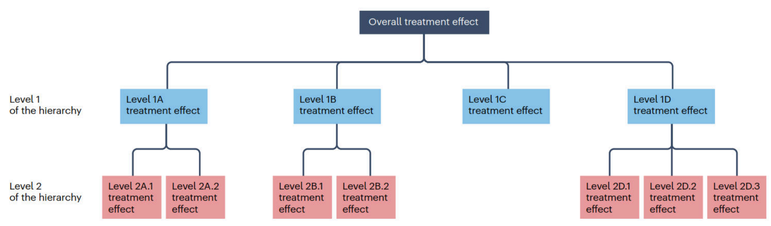
- Bayesian analysis may make a study "smaller – but more complex and expensive", while creating efficiencies in the entire clinical development program
- Vague incentive for drug developers
- Safety data maybe cost driver not efficacy data

## Recommended Methods

**Bayesian hierarchical models (BHM)**

- assumes a common treatment effect
- allows "borrowing information" between subgroups

**Figure 1.** Hierarchical Model (Ruberg et al. 2023)




**Equation 1.** Bayesian Hierarchical Model (Chu & Yuan, 2018)

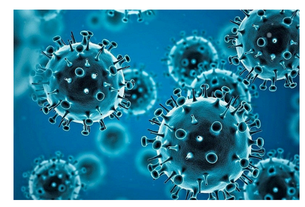
$$y_j|p_j \sim Bin(p_j)$$
$$\theta_j = \log\left(\frac{p_j}{1-p_j}\right)$$
$$\theta_j|\theta,\sigma^2 \sim N(\theta,\sigma^2)$$
$$\theta \sim N(\alpha_0,\omega_0^2)$$

- Issues:
  - If objective is to detect differential treatment effect, assumption of a common treatment effect not plausible
  - original BHM proposed by Berry et al.(2013) can show an inflation of the nominal risk of false positive conclusions of 10% to over 50% (Chu & Yuan, 2018)
- Solution:
  - Bayesian predictive cross-validation models (Dias et al, 2011) predict treatment effect in subgroup of interest based on shared treatment effect of all other (!) subgroups

## Empirical Studies

**Pfizer/BioNTech COVID-19 vaccine study**

- Pfizer/BioNTech evaluated the vaccine efficacy of the BNT162b2 ("Comirnaty") mRNA vaccine using Bayesian statistics



- Bayesian statistics did not add value because of very strong effects
- Questionable assumptions, e.g., prior distribution for vaccine efficacy centered around 30% labelled as "pessimistic" and "minimally informative"

**Pediatric Cardiac Arrest Trial (THAPCA-OH)**

- Primary Frequentist analysis: Therapeutic hypothermia rejected as ineffective treatment after not reaching the standard statistical significance threshold of 5%
  → ineffective treatment


- Secondary Bayesian analysis: Therapeutic hypothermia had a 94% probability of any benefit over therapeutic normothermia
  → effective treatment

## Key Solutions

Rethink the line of argumentation:

- Balance the presentation of differences and (!) similarities


- Target communication to audience, esp. regulators
- Evaluate operating characteristics based on regulatory standards, esp. risk for false positive conclusions
- Promote Bayesian studies which show added value, esp. for patients

**References**

1. Campbell, G. (2020). FDA Regulatory Acceptance of Bayesian statistics. In: Lesaffre, E., Baio, G., & Boulanger, B. Bayesian Methods in Pharmaceutical Research. Boca Raton, FL, USA: CRC Press.
2. Ruberg, S. J. et al. Application of Bayesian approaches in drug development: starting a virtuous cycle. Nat Rev Drug Discov 1–16 (2023) doi:10.1038/s41573-023-00638-0.
3. Igl, W. A review of the "Virtuous Cycle" paper on Bayesian statistics by Ruberg et al. (2023). https://wilmarigl.de/?p=689 (2023).
4. Gill, J. Bayesian Methods: A Social and Behavioral Sciences Approach. (Chapman and Hall/CRC, 2007).
5. Rubin, D. B. For objective causal inference, design trumps analysis. The Annals of Applied Statistics 2, 808–840 (2008).
6. Polack, F. P. et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. New England Journal of Medicine 383, 2603–2615 (2020).
7. Senn, S. The design and analysis of vaccine trials for COVID-19 for the purpose of estimating efficacy. Pharmaceutical Statistics 21, 790–807 (2022).
8. Harhay, M. O. et al. A Bayesian Interpretation of a Pediatric Cardiac Arrest Trial (THAPCA-OH). NEJM Evidence 2, EVIDoa2200196 (2022).
9. Chu, Y. & Yuan, Y. A Bayesian basket trial design using a calibrated Bayesian hierarchical model. Clin Trials 15, 149–158 (2018).
10. Dias, S., Sutton, A. J., Welton, N. J. & Ades, A. E. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. http://www.nicedsu.org.uk (2011).
11. McElreath, R. Statistical Rethinking - A Bayesian Course with Examples in R and STAN. (Chapman & Hall, 2020).