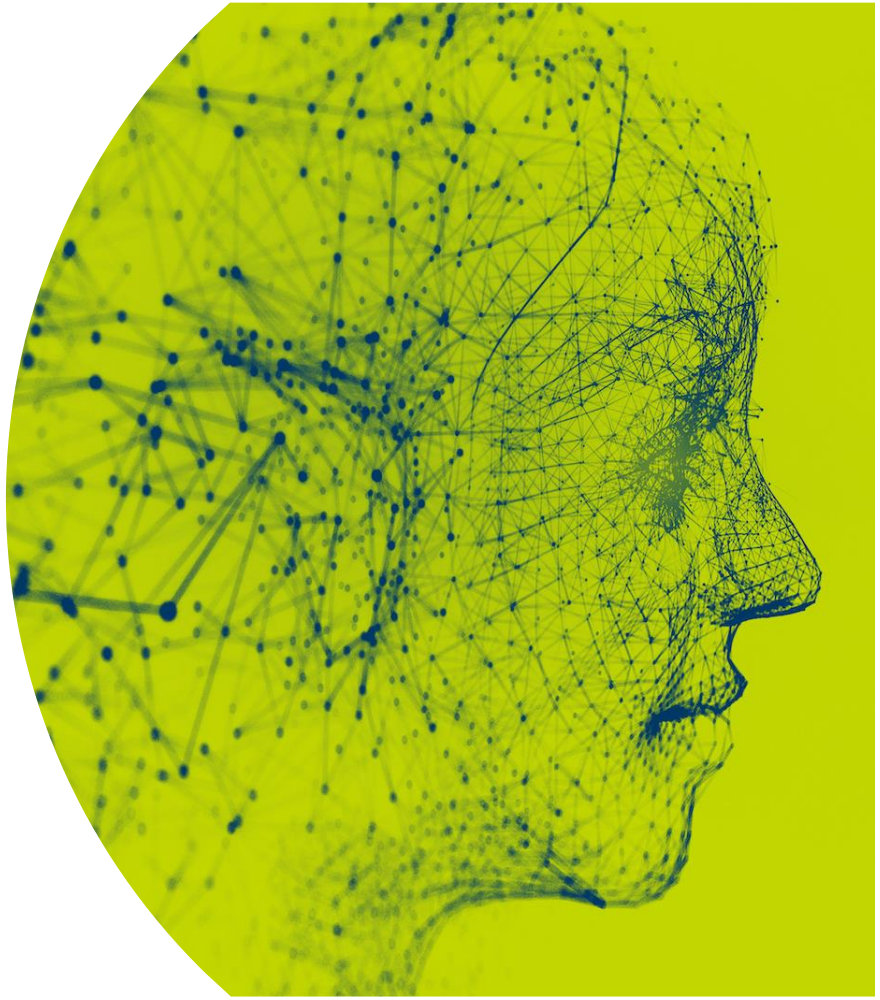
AstraZeneca

A conservative approach to leveraging external evidence for effective clinical trial design

https://arxiv.org/abs/2211.02381



Fabio Rigat*

Speaker: Binbing Yu**

October 2022

* fabio.rigat@astrazeneca.com

** binbing.yu@astrazeneca.com

Bayesian clinical trial design is work in progress

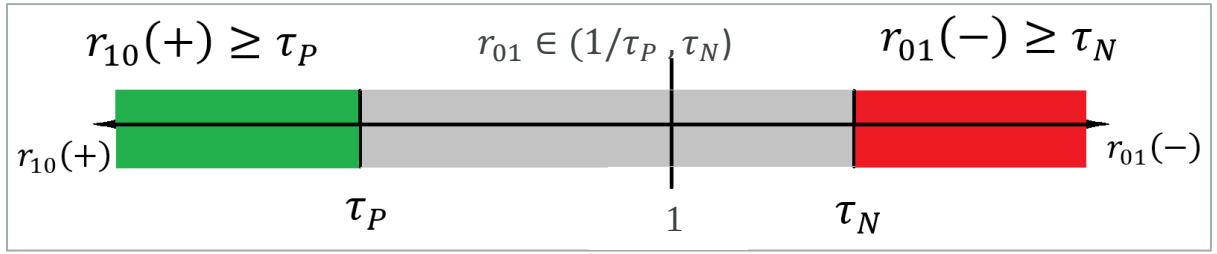
- **Any trial design requires grounding in background evidence**
  - Use of prior evidence is not a unique to Bayesian methods
- **Bayesian trial design is currently mostly applied to:**
  - calculating assurance, to mitigate reliance on narrow design assumptions [4-5],
  - borrowing/extrapolation: dynamic modeling of study outcomes [6-8],
  - use of external controls [9]
- **Bayesian trial design methodology is work in process**
  - due to lack of *transparent* and *robust* prior estimation process/guidance.
- **We propose a simple Bayesian method** providing
  - confidence to "skeptical consumers of statistics" [10], ensuring designs *against* priors,
  - a pragmatic and ethically compelling base for clinicians, addressing equipoise dilemmas

https://www.clipartwiki.com

Posterior odds of clinical hypotheses lead to effective designs

- **Trials are needed when background evidence is insufficient**
  - confirmatory designs can provide ***conclusive overall evidence*** (background + trial outcome), adequate to *disturb an initial clinical equipoise* [11-12] beyond pre-specified positive ***and*** negative evidence thresholds (τ_P , τ_N).
- **Bayesian solution:** "Bayesian characteristics" (BACs) i.e. odds of design hypotheses
  - "+"/"−": a positive/negative study (e.g. p-val </≥ 0.05, p(efficacy gain > δ| data, priors) >/≤ 0.95),
  - design operational characteristics (OCs): $p(-|H_0), p(+|H_1)$ i.e. specificity and sensitivity of the primary trial outcome,
  - pre-study odds of H_0 vs H_1: $r_{01} = \frac{P(H_0)}{P(H_1)}$ or $r_{10} := 1/r_{01} = \frac{P(H_1)}{P(H_0)}$, estimated from background evidence
  - post-study odds of H_0 vs H_1:
    $$\begin{cases} r_{01}(-) = \frac{P(H_0|-)}{P(H_1|-)} = r_{01} \times LR(-) = r_{01} \times \frac{p(-|H_0)}{1-p(+|H_1)} & \text{should the trial be negative,} \\ r_{10}(+) = \frac{P(H_1|+)}{P(H_0|+)} = r_{10} \times LR(+) = r_{10} \times \frac{p(+|H_1)}{1-p(-|H_0)} & \text{should the trial be positive.} \end{cases}$$
  - **BACs design**: sample size s.t. $r_{01}(-) \geq \tau_N > \max(1, r_{01})$ ***and*** $r_{10}(+) \geq \tau_P > \max(1, r_{10})$.

Strong positive evidence



Strong negative evidence

BACs design shows strength of positive and negative evidence

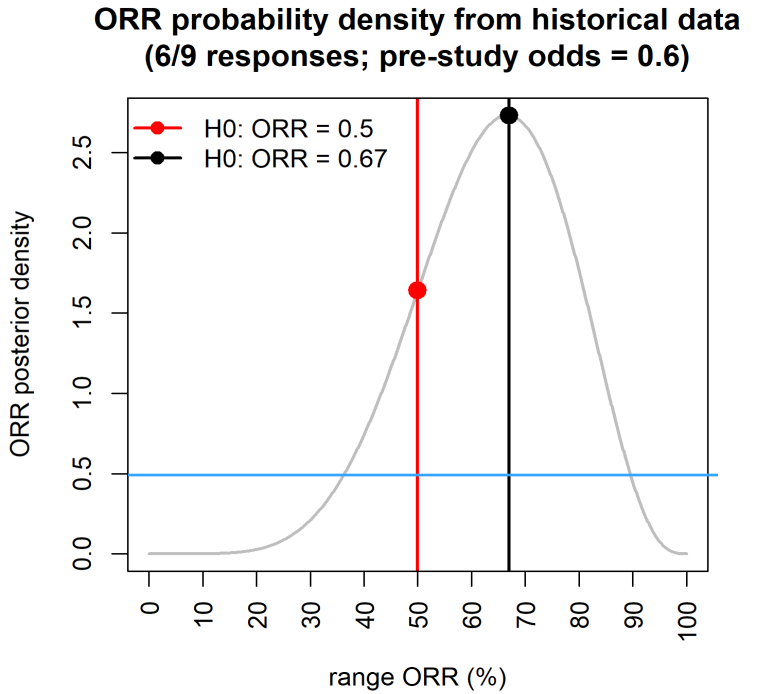
- **Worthless designs** (coin flips) carry no information about the truth of the design hypotheses
- **Current OCs for confirmatory trials** turn 1:1 pre-study odds into
  - $\tau_N \approx$ 5-10 odds in favor of H_0 if trial is negative,
  - $\tau_P \approx$ 16-18 odds in favor of H_1 if trial is positive
- **OCs symmetry implies BACs symmetry** when pre-study odds are 1:1
- **Optimistic design priors (2:1 odds in favor of H_1) weaken negative evidence:** need greater N to achieve strong negative evidence
- **Pessimistic design priors (2:1 odds in favor of H_0) weaken positive evidence:** need greater N to achieve strong positive evidence

Prior evidence r_{01}	Primary trial outcome operating characteristics		BACs	
	Specificity	Sensitivity	$r_{01}(-)$	$r_{10}(+)$
any value	50%	50%	r_{01}	$1/r_{01}$
1	95%	80%	4.75	16
		90%	9.5	18
	90%	90%	9	9
	80%	80%	4	4
1/2	95%	80%	2.38	32
	90%	90%	4.5	18
	80%	80%	2	8
2	95%	80%	9.5	8
	90%	90%	18	4.5
	80%	80%	8	2

BACs design – testing for one proportion

- **Ph1 study:** ORR% = 6/9 ≈ 67% in evaluable subjects:
  - $\boldsymbol{H_1}$ = 67% is the mode of the ORR posterior distribution *Beta(1 + 6, 1+ 9 – 6)*
  - $\boldsymbol{H_0}$ = 50% is the SOC ORR% in this population (more background evidence!)
  - r_{01} = **dbeta(0.5,7,4)/dbeta(0.67,7,4)** ≈ 0.6 i.e. pre-study odds ≈ 3:2 in favor of H_1
- **OCs of single arm ORR confirmatory study** (Ph1 expansion/ Ph1B /Ph2),
  - Specificity: $p(+|H_0) := 1 - p(-|H_0)$ = 1 - pbeta(NR,N,p = $\boldsymbol{H_0}$) > 0.95 where N is sample size and NR ∈ {0, *N*} is n. responders
  - Sensitivity: $p(+|H_1) :=$ power.prop.test(N, p1 = $\boldsymbol{H_0}$, p2 = $\boldsymbol{H_1}$)
- **Ph2 single arm study sample size:**
  - **N=140** carries ≈ 80% power and strong BACs if 1:1 pre-study odds are used (base case)
  - **N=170** carries ≈ 90% power and *stronger* BACs than a typical confirmatory study using 3:2 pre-study odds (Ph1-estimated design prior)
- **BACs cost of ensuring Ph2 go/no-go against a possibly overoptimistic design prior:**

  (170-140)/140 ≈ 21% increase in N.



N	Primary trial outcome operating characteristics		BACs			
			$r_{01} = 0.6$		$r_{01} = 1$	
	Specificity	Sensitivity	$r_{01}(-)$	$r_{10}(+)$	$r_{01}(-)$	$r_{10}(+)$
130	95%	79.8%	2.8	26.6	4.7	16.0
140		82.7%	3.3	27.5	5.5	16.5
150		85.2%	3.9	28.4	6.4	17.0
160		87.4%	4.5	29.1	7.5	17.5
170		89.2%	5.3	29.7	8.8	17.9
180		90.9%	6.3	30.3	10.4	18.2

BACs summary & some open questions

- **BACs are not a methodological development** but
  - a simple, well-understood Bayesian approach to study design seamlessly incorporating frequentist OCs and clearly identifying what is needed from background evidence, i.e. a single additional input (r_{01}) compared to current practice,
  - an ethical approach requiring study design to be commensurate to initial level of clinical equipoise,
  - a skeptical use of design priors, where study sample size can only increase in the amount of prior information.
- **"Don't run before you can walk":** robust BACs design implementation requires guidance ensuring that background evidence used for trial design is
  - represented clearly and fairly (principles for source selection) and
  - synthesized appropriately (principles for data integration / r_{01} estimation)
- **Should dynamic prior discounting be applied to adaptive BACs design?** Would provide a less conservative model, e.g.
  - start from full insurance against the design prior and
  - allow for a pre-specified level of trust building along the study, *iff* design prior and IA study outcome agree

Thank you for your feedback!

Acknowledgements

Daniel Jackson, Michael Sweeting, Peter De Porre, Stephen Weng, Telba Irony, Vladimir Dragalin, Tobias Mielke, Dwaine Banton, Filip De Ridder, Huiming Liao, Rafael Sauter, Maoji Li, Fowzia Ibrahim, Professor David Spiegelhalter (University of Cambridge), Professor Jim Smith (University of Warwick)

References

1. EMA 2017 Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products
2. FDA 2005 Guideline on Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers
3. FDA Project Optimus
4. O'Hagan A, Stevens JW, Campbell MJ. Assurance in clinical trial design. Pharmaceutical Statistics. 2005;4:187-201
5. Crisp, A, Miller, S, Thompson, D, Best, N. Practical experiences of adopting assurance as a quantitative framework to support decision making in drug development. Pharmaceutical Statistics. 2018; 17: 317– 328
6. Neuenschwander, B., Branson, M. and Spiegelhalter, D.J. (2009), A note on the power prior. Statist. Med., 28: 3562-3566
7. Hobbs, B.P., Carlin, B.P., Mandrekar, S.J. and Sargent, D.J. (2011), Hierarchical Commensurate and Power Prior Models for Adaptive Incorporation of Historical Information in Clinical Trials. Biometrics, 67: 1047-1056
8. Louis Dron, Shirin Golchi, Grace Hsu, Kristian Thorlund (2019) Minimizing control group allocation in randomized trials using dynamic borrowing of external control data – An application to second line therapy for non-small cell lung cancer. Contemporary Clinical Trials Communications, 16:100446
9. Mahta Jahanshahi, Keith Gregg, Gillian Davis, Adora Ndu, Veronica Miller, Jerry Vockley, Cecile Ollivier, Tanja Franolic, Sharon Sakai (2021) The Use of External Controls in FDA Regulatory Decision Making. Ther Innov Regul Sci. 2021; 55(5): 1019–1035.
10. The Statistics Wars and Their Casualties
11. Benjamin Freedman (1987), Equipoise and the ethics of clinical research, N. Engl.J.Med, 317:141-145
12. Franklin G. Miller and Steven Joffe (2011) Equipoise and the dilemma of randomized clinical trials, N. Engl.J.Med, 364:476-480
13. Nicholas Johnson, Richard J Lilford, Wayne Brazier (1991), At what level of collective equipoise does a clinical trial become ethical? Journal of Medical Ethics, 17:30-34