Core Clinical Sciences

# Adaptive Trial Designs for Public and Global Health

## *Vaccine and Cluster Randomized Trials go Bayesian*

Ofir Harari ofir@coreclinicalsciences.com

Adaptive Platform Trial Scientific Meeting, Toronto, ON

September 28, 2023

# Background

# Global and Public Health Trial Characteristics

- Typically large
- Are run for a long time
- Multi-center, logistically challenging
- Are funded by the taxpayer or nonprofits
- Potentially high impact
- Could benefit greatly from saving time and resources

# Bayesian Adaptive Trial Designs

- Are simulation-guided, and are therefore flexible – but also computer-intensive.
- Consists of multiple arms and multiple stages with early arm-stopping options.
- The decision boundaries are set for the *posterior probability of efficacy*

$$p_i^{\text{effic}} = \mathbb{P}\left(\text{Treatment } i \text{ is efficacious} \middle| \text{data}\right)$$

  for trials testing treatments against a common control, or the *posterior probability of superiority*

$$p_i^{\text{sup}} = \mathbb{P}\left(\text{Treatment } i \text{ is the most efficacious} \middle| \text{data}\right)$$

  for trials comparing all therapies against each other. These get updated at every interim readout.
- Potential gains include earlier EUA application, fewer patients receiving futile treatments, fewer overall patients treated and shorter run times.

# Multi-arm, Multi-stage Vaccine Trials

# Modelling Assumptions

- All arms' case counts follow Poisson processes.
- *Vaccine effects* (VEs) are measured in terms of the rate reduction relative to control.
- Once a certain total number of cases has been recorded, data are unblinded and case distribution across groups is analyzed.
- Some transformation of the VEs can be shown to follow a multinomial distribution.
- Using a conjugate Dirichlet prior distribution, the posterior distribution is also Dirichlet.

# The Main Idea

- If we define

$$\theta_i = \frac{q_i(1-\mathrm{VE}_i)}{q_{\mathrm{ctrl}} + \sum_{j=1}^{K} q_j(1-\mathrm{VE}_j)},$$

where $q_i$ is the allocation probability of the $i^{\mathrm{th}}$ vaccine, and assign $\boldsymbol{\theta}$ a Dirichlet prior, then

$$\boldsymbol{\theta} \big| n_1(t), \ldots n_K(t), n_{\mathrm{ctrl}}(t)$$

will also follow a Dirichlet distribution, where $n_i(t)$ denotes the number of infections recorded in the $i^{\mathrm{th}}$ vaccine group.

- Inference on the VEs is achievable via the back-tranform

$$\mathrm{VE}_i = 1 - \frac{\theta_i}{\theta_{\mathrm{ctrl}}} \frac{q_{\mathrm{ctrl}}}{q_i}.$$

# What We Did

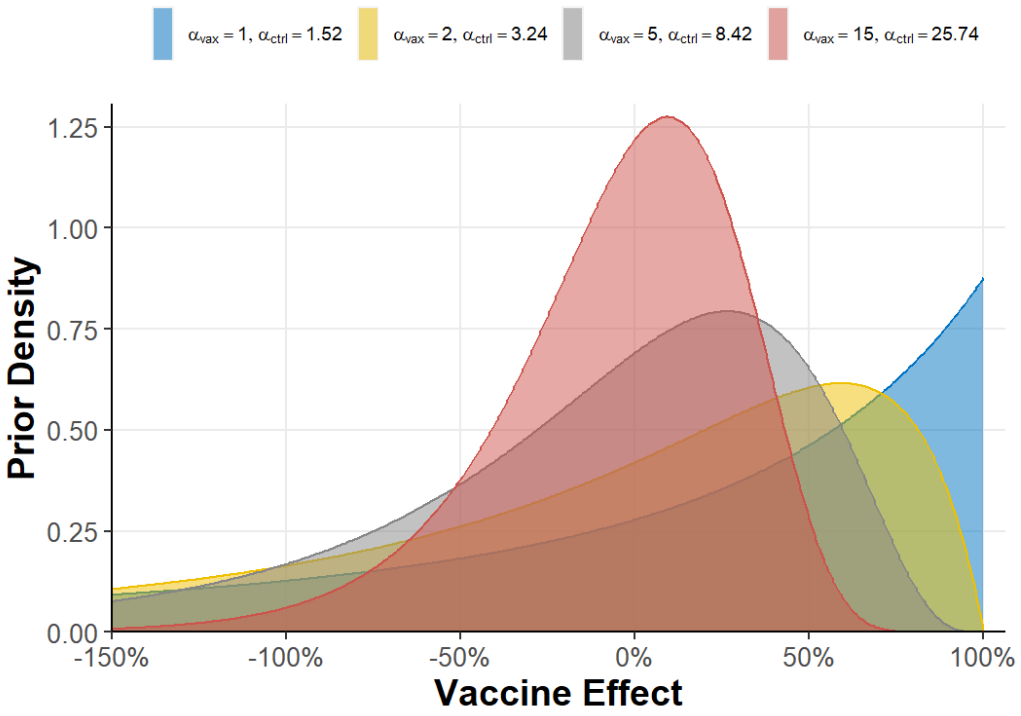
- Generalized the model used by Pfizer/BioNTech to any number of arms with any allocation ratio.
- Calculated the joint VE distribution and showed that inference can be restricted to paired vaccine-placebo comparisons.
- Proposed a principled way for choosing a prior distribution.
- Derived closed form expressions for all inferential quantities, including $p^{\text{effic}}$ .
- We can now run 100,000 simulations in ~1 minute on a standard PC.

# Trial Design Example

- Three vaccine groups and a common control (placebo).
- A $\sqrt{3} : 1 : 1 : 1$ Dunnett allocation ratio to minimize the average VE uncertainty.
- We assume VEs of 50%, 37.5% and 25%.
- Interim analysis take place when an average of 10.5 events are observed across the remaining arms.
- Early stopping due to efficacy occurs if $p^{\text{effic}} > 99.025\%$ at any analysis.
- Early stopping due to futility occurs if $p^{\text{effic}} < 25\%, 50\%$ or $75\%$ in the first, second or third analysis, respectively.

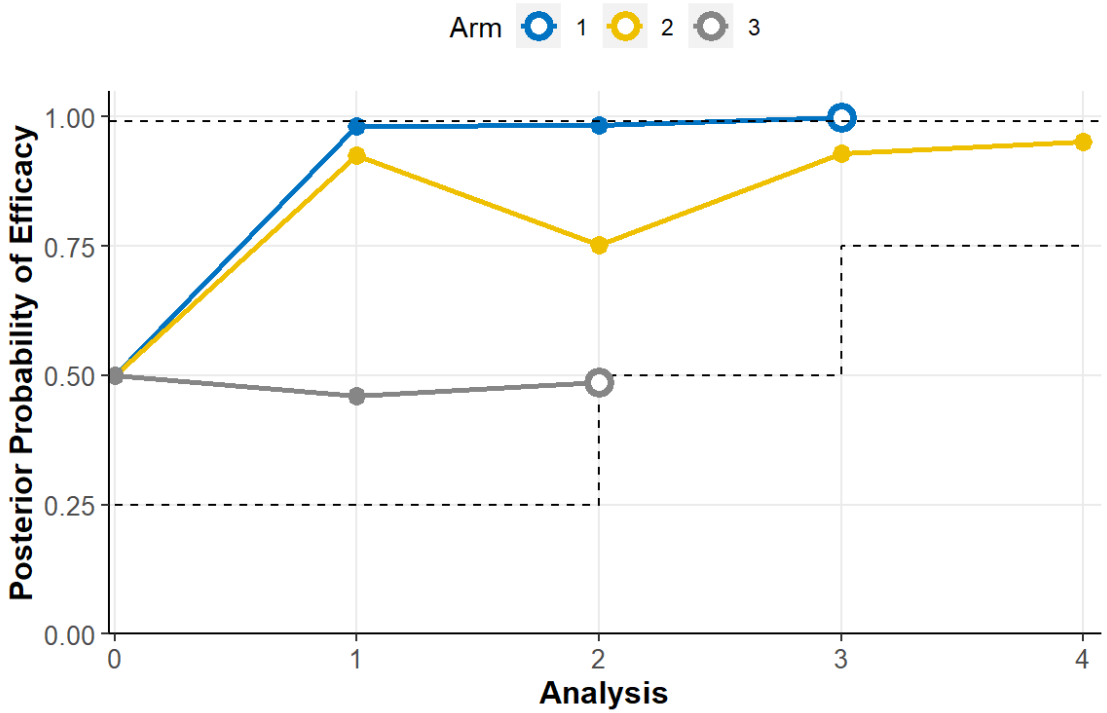
# Prior Selection

- Prior candidates are required to have a prior probability of efficacy of $50\%$.
- We compare multiple candidates and use summaries and density plots.




| $\alpha_{\text{vax}}$ | $\alpha_{\text{ctrl}}$ | Mean | Median | 95% CrI |
|---|---|---|---|---|
| 1.00 | 1.52 | -2.33 | 0.00 | [-16.85; 0.97] |
| 2.00 | 3.24 | -0.55 | 0.00 | [-5.27; 0.88] |
| 5.00 | 8.42 | -0.17 | 0.00 | [-2.01; 0.70] |
| 15.00 | 25.74 | -0.05 | 0.00 | [-0.87; 0.49] |

# A Sample Trial

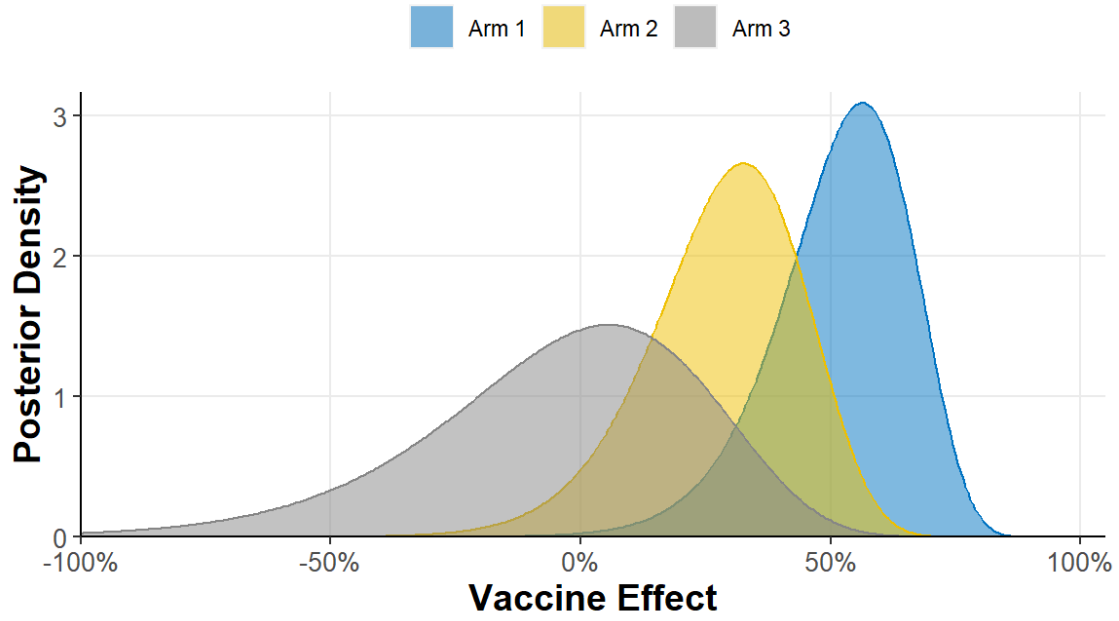



- In this randomly simulated trial, the third vaccine arm was stopped for futility after two analyses
- The first vaccine was found to be efficacious after three analyses.
- The second vaccine arm completed four analyses without meeting the efficacy threshold.

# Posterior Inference

- Point and interval estimation is based on the posterior VE distribution –

| Arm | Stopping reason | Analysis 1 cases | Analysis 2 cases | Analysis 3 cases | Analysis 4 cases | Efficacy | Pr(Efficacy\|Data) | VE [95% CrI] |
|---|---|---|---|---|---|---|---|---|
| Control | | 20 | 36 | 57 | 72 | | | |
| 1 | Early efficacy | 4 | 10 | 15 | | Yes | 99.8% | 53.4% [20.7%; 74.2%] |
| 2 | End of trial | 6 | 17 | 23 | 29 | No | 95.2% | 29.8% [-6.3%; 54.8%] |
| 3 | Early futility | 12 | 21 | | | No | 48.6% | -0.9% [-69.7%; 41.4%] |



# Operating Characteristics (100k simulations)

| Arm | VE | Power | $\mathbb{P}(\text{Stop} = 1)$ | $\mathbb{P}(\text{Stop} \leq 2)$ | $\mathbb{P}(\text{Stop} \leq 3)$ |
|---|---|---|---|---|---|
| All | 0% | 2.5% | 23.9% | 53.2% | 79.0% |
| 1 | 50% | 81.3% | 17.8% | 43.1% | 70.9% |
| 2 | 37.5% | 50.1% | 10.8% | 25.9% | 50.4% |
| 3 | 25% | 22.8% | 10.3% | 24.1% | 48.2% |

- Trial inputs were calibrated for a $2.5\%$ type I error rate and $\geq 80\%$ power for a vaccine with $50\%$ VE.
- A high probability of early stopping – but not necessarily fewer doses administered.

# Adaptive Cluster Randomized Trials

# Some CRT Background

- Trials in which the cluster (e.g., a village) is the randomization unit.
- As an example: laying attractive toxic sugar baits outside the village to reduce malaria infections.
- Induces two variance components: subject-to-subject and cluster-to-cluster.
- The variance inflation is governed by the cluster size and the *intrucluster correlation coefficient* (ICC).
- A large ICC or cluster size can significantly inflate the sample size.

# What We Did

- Proposed the Poisson-Gamma mixture model for CRTs with binary endpoints.
- Showed that some transformation of the mean event rates has conjugate beta prior distributions.
- Established the mathematical relation between the Gamma random effect parameters and the ICC.
- Proposed a principled way for choosing prior distributions.
- Derived the empirical Bayes plug-in estimators for the Gamma shape parameters.
- We can now run 100,000 simulations in ~10 minutes on a standard PC.

# Modelling Assumptions

- For reasonably large clusters of size $n_i$, and a fairly low event rate $r_i$, we can model

$$Y_{ij}|r_{ij} \sim \text{Poisson}\left(n_i r_{ij}\right)$$

the number of cases observed in the $j^{\text{th}}$ cluster or the $i^{\text{th}}$ arm.

- To capture cluster-to-cluster variability, we model

$$r_{ij}|\mu_i, \phi_i \sim \text{Gamma}\left(\phi_i, \mu_i/\phi_i\right),$$

where $\mu_i$ is now the mean event rate and $\phi_i$ a shape parameter.

- The marginal case number distribution can be shown to be negative binomial with

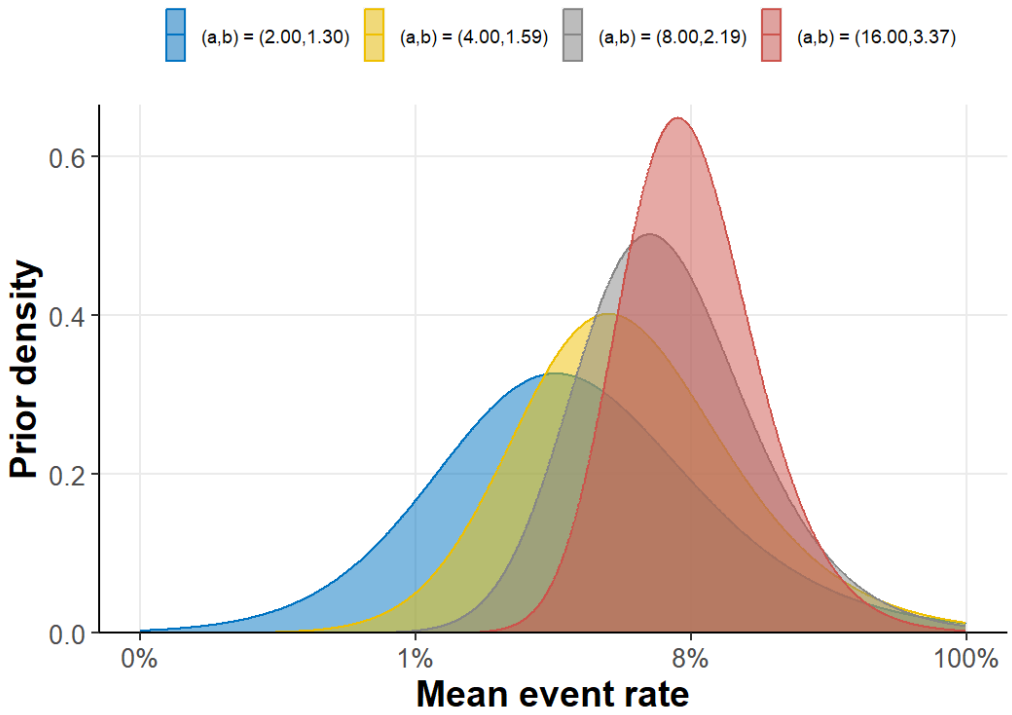
$$p_i = \frac{\mu_i n_i}{\phi_i + \mu_i n_i},$$

and has a conjugate Beta prior.

# Prior Selection

- All priors are subject to satisfying a pre-specified a priori mean event rate (in this case $10\%$), assuming an ICC (here $0.15$ and $0.2$).

- We then compare them by their summaries.

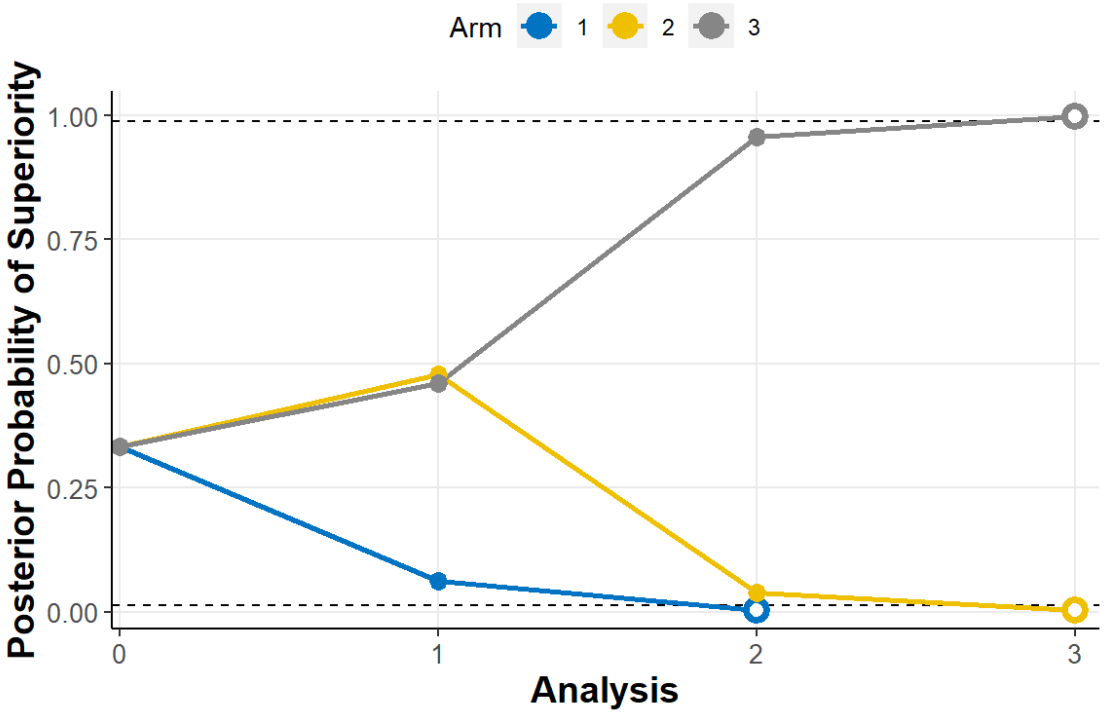



| $\mathbb{E}[\mu]$ | $\rho$ | $\phi$ | $(a, b)$ | 95% CrI |
|---|---|---|---|---|
| 0.10 | 0.15 | 0.74 | $(2.00, 1.30)$ | [0.22%; 46.29%] |
| | | | $(4.00, 1.59)$ | [0.70%; 48.32%] |
| | | | $(8.00, 2.19)$ | [1.49%; 41.25%] |
| | | | $(16.00, 3.37)$ | [2.56%; 31.67%] |
| | 0.20 | 0.56 | $(2.00, 1.22)$ | [0.18%; 42.06%] |
| | | | $(4.00, 1.44)$ | [0.56%; 48.23%] |
| | | | $(8.00, 1.89)$ | [1.24%; 44.36%] |
| | | | $(16.00, 2.78)$ | [2.21%; 35.14%] |

# Adaptive Design Example

- Suppose that we wish to trial three treatments
- Assuming mean event rates of $10\%$, $7.5\%$ and $5\%$
- A maximum number of $550$ clusters divided into three analyses
- Average cluster size of $50$,
- An arm gets stopped for futility if $p^{\text{sup}} \leq 1.2\%$
- The trial gets stopped if any arm achieves $p^{\text{sup}} > 98.8\%$
- We simulate the data with an ICC of $0.15$

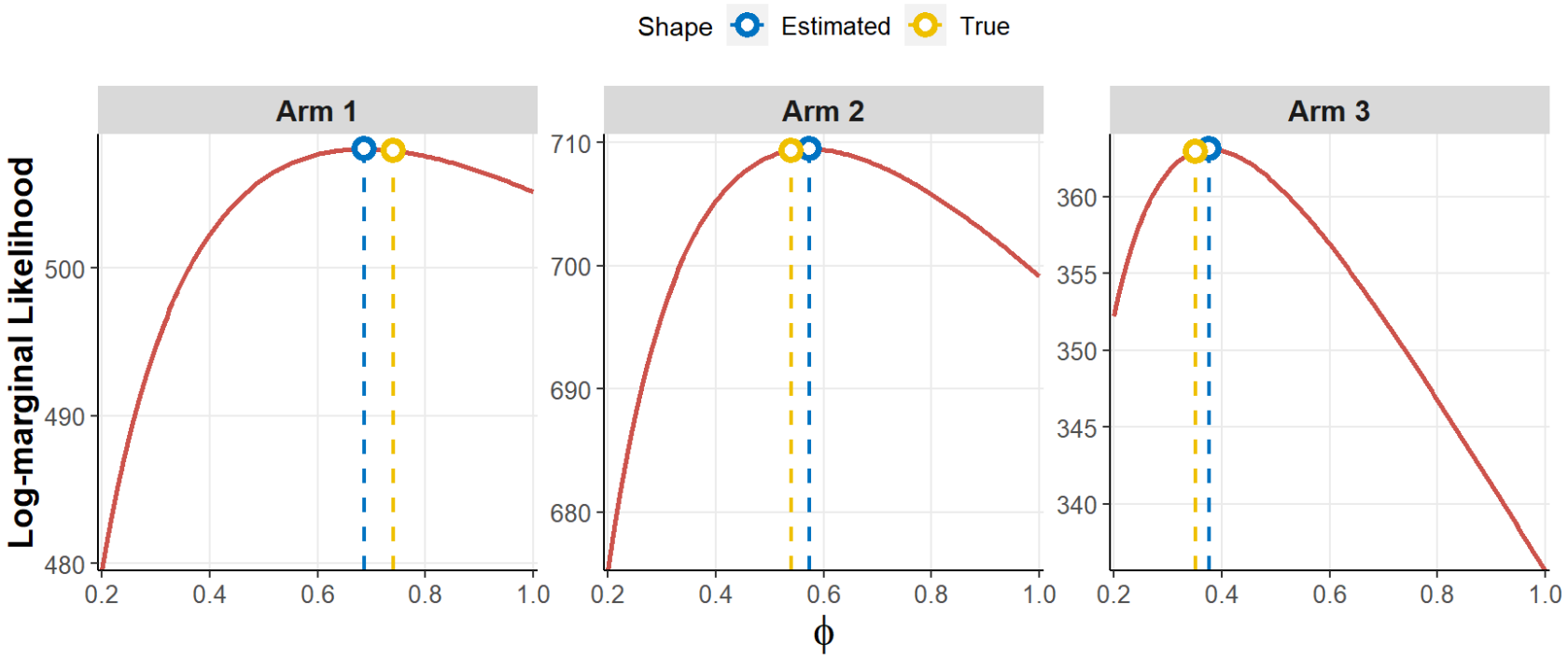
# A Sample Trial




- Here first arm was dropped after the second analysis for futility.
- The third treatment was found to be the superior of the three after the third analysis.

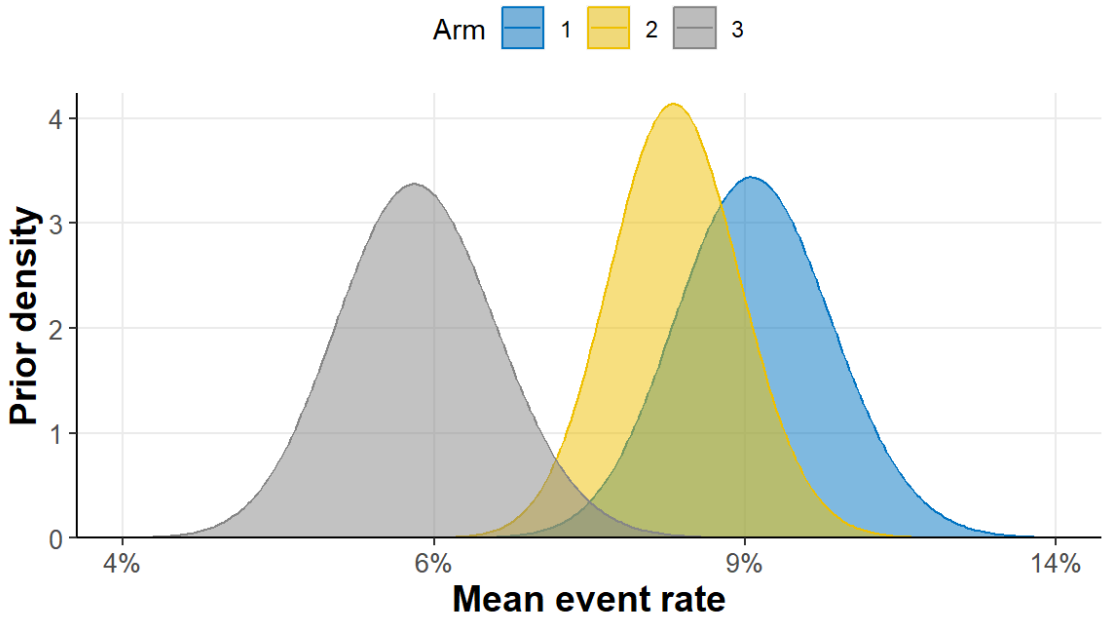
# Shape Parameter Estimation via Empirical Bayes

- The negative binomial model is conditional on knowing the values of the shape parameters $\phi_i$.
- As plug-in estimates, we propose maximizing the log-marginal likelihood components.



# Posterior Estimation and Inference

| Trt | Analysis 1<br>Clusters, Cases/Patients | Analysis 2<br>Clusters, Cases/Patients | Analysis 3<br>Clusters, Cases/Patients | Outcome | Mean event rate<br>[95% CrI] |
|---|---|---|---|---|---|
| 1 | 62, 261/3075 | 123, 559/6203 | | Early inferiority | 8.9% [7.2%; 11.3%] |
| 2 | 62, 202/3057 | 123, 469/6022 | 213, 844/10558 | | 7.9% [6.6%; 9.6%] |
| 3 | 62, 206/3059 | 123, 341/6186 | 213, 582/10652 | Superiority | 5.4% [4.3%; 6.9%] |



# Operating Characteristics (100K Simulations)

| Mean event rates | Mean no. of clusters | % of patients in Arm 1 | % of patients in Arm 2 | % of patients in Arm 3 | True positive rate | False positive rate |
|---|---|---|---|---|---|---|
| (10.0%, 7.5%, 5.0%) | 412.1 | 20.9% | 38.8% | 40.3% | 80.6% | 0.0% |
| (10.0%, 10.0%, 10.0%) | 545.7 | 33.3% | 33.3% | 33.3% | – | 2.0% |
| (10.0%, 5.0%, 5.0%) | 532.0 | 13.0% | 43.5% | 43.5% | – | 9.7% |

- Note that the “null hypothesis” here includes every scenario where there is no clear-cut winner.
- “False positives” include detection of one of two equally effective treatments.
- For a “true positive”, we must not only detect difference but also declare the correct winner.
- In a standard trial, the 1-2 comparison alone at the 5% level (one-sided) would require 528 clusters.
- Also note the proportion of patients receiving the worst treatment.

# Concluding Remarks

# Famous Last Words

- Global health trials could use every bit of efficiency.
- Adaptive trial designs provide just that – but require many simulation runs.
- We propose innovative, MCMC-free, Bayesian vaccine and cluster randomized trial designs.
- Response adaptive randomization (RAR) is unlikely to be applicable here.
- Consider time to outcome and enrollment rate: could your trial benefit from adaptations?
- Possible extensions: time-varying VE; stepped-wedge CRTs.
- Ask us for the preprint!

**Thank you!**

# Bibliography

# References


1. Hayes, R. J. & Moulton, L. H. *Cluster randomised trials*. (Chapman; Hall/CRC, 2017).

2. Laurent, Stéphane & Legrand, Catherine. A bayesian framework for the ratio of two poisson rates in the context of vaccine efficacy trials. *ESAIM: PS* **16**, 375–398 (2012).

3. Senn, S. The design and analysis of vaccine trials for COVID-19 for the purpose of estimating efficacy. *Pharmaceutical Statistics* **21**, 790–807 (2022).

4. Berry, S. M., Carlin, B. P., Lee, J. J. & Muller, P. *Bayesian adaptive methods for clinical trials*. (CRC Press, 2010).

5. Food and Drug Administration. Adaptive designs for clinical trials of drugs and biologics: Guidance for industry. (2019).

6. Turner, R. M., Omar, R. Z. & Thompson, S. G. Bayesian methods of analysis for cluster randomized trials with binary outcome data. *Statistics in Medicine* **20**, 453–472 (2001).

7. Shen, J., Golchi, S., Moodie, E. E. & Benrimoh, D. Bayesian group sequential designs for cluster-randomized trials. *Stat* **11**, e487 (2022).

8. Casella, G. An introduction to empirical bayes data analysis. *The American Statistician* **39**, 83–87 (1985).