MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Bayesian Methods for Clinical Trials

Lecture 7: Basket Clinical Trials

Libby Daniells & Pavel Mozgunov & Thomas Jaki

MRC Biostatistics Unit

February 25, 2025

Master protocols

Clinical trials are most commonly based on ‘one disease, one treatment and one population’

Master protocols are a *new* type of study that seeks to answer multiple questions within a single study

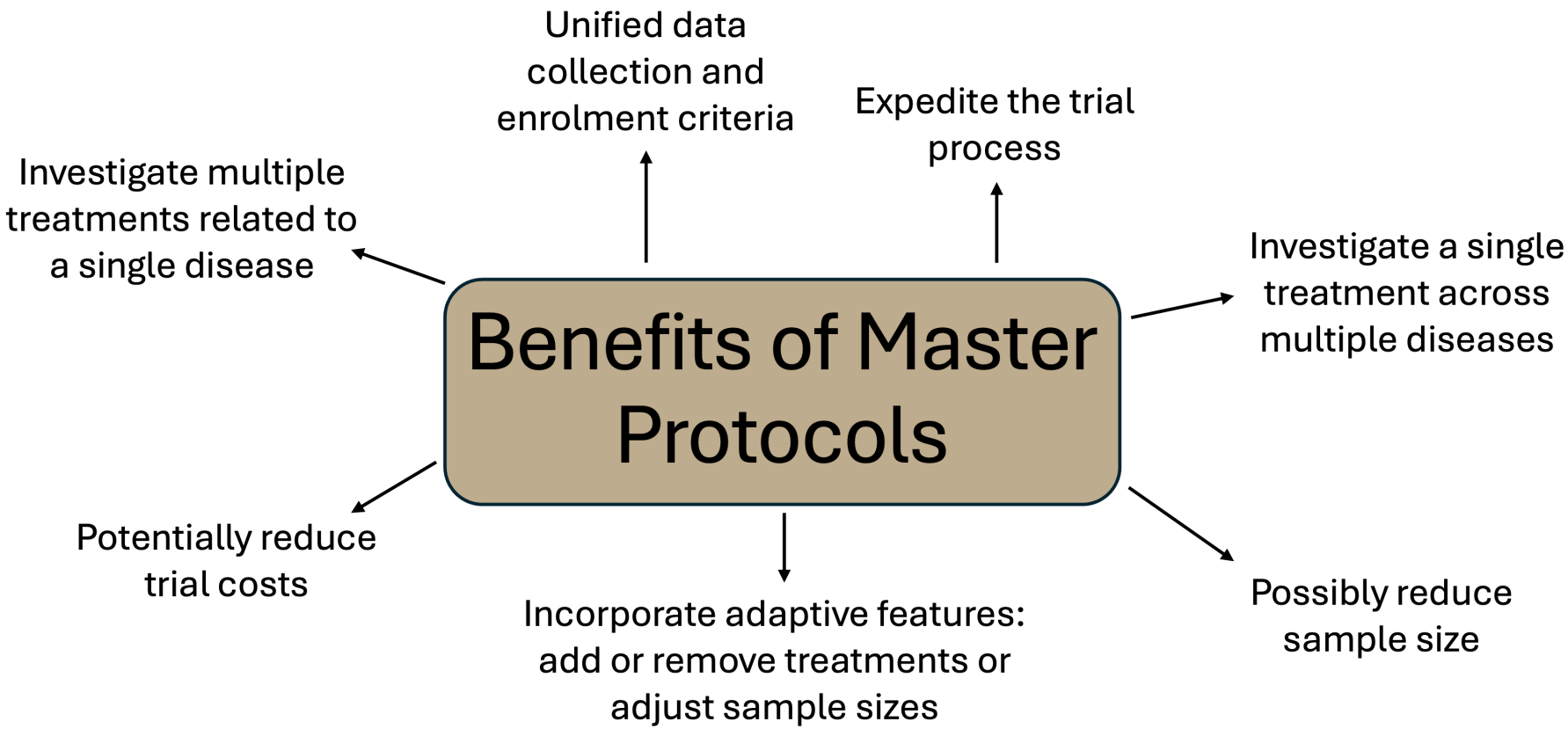
- Platform trials
- Basket trials
- Umbrella trials

Driven by the feasibility of **precision medicine**.

Precision Medicine

- ‘One-size fits all’ often does not apply when targeting a treatment to patients with the same disease.
- Molecular profiling and genetic testing at the individual patient level has become more feasible and affordable.
- **Biomarker**: a measurable indicator of biological properties or genetic aberration.
- In precision medicine treatments are targeted to an individuals biomarkers, as opposed to a disease type on a whole. Tailors treatments to a patients intrinsic factors.
- **Key objective**: increased efficiency for drug development.

Benefits

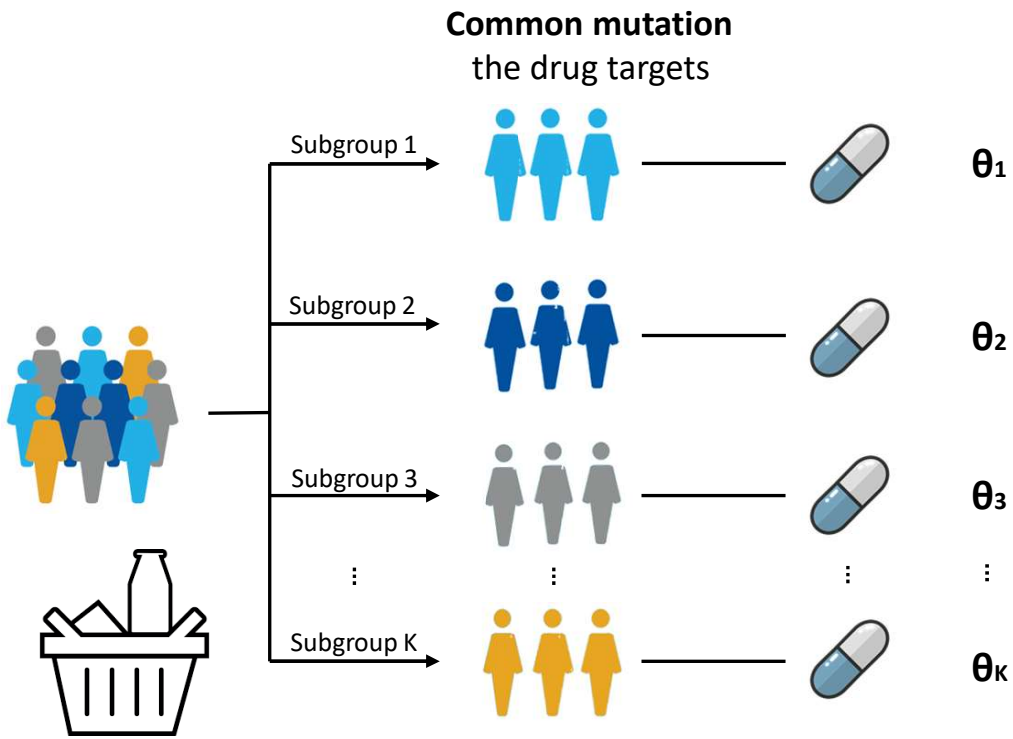

Benefits of Master Protocols
Unified data collection and enrolment criteria
Expedite the trial process
Investigate multiple treatments related to a single disease
Investigate a single treatment across multiple diseases
Potentially reduce trial costs
Incorporate adaptive features: add or remove treatments or adjust sample sizes
Possibly reduce sample size

Basket trials

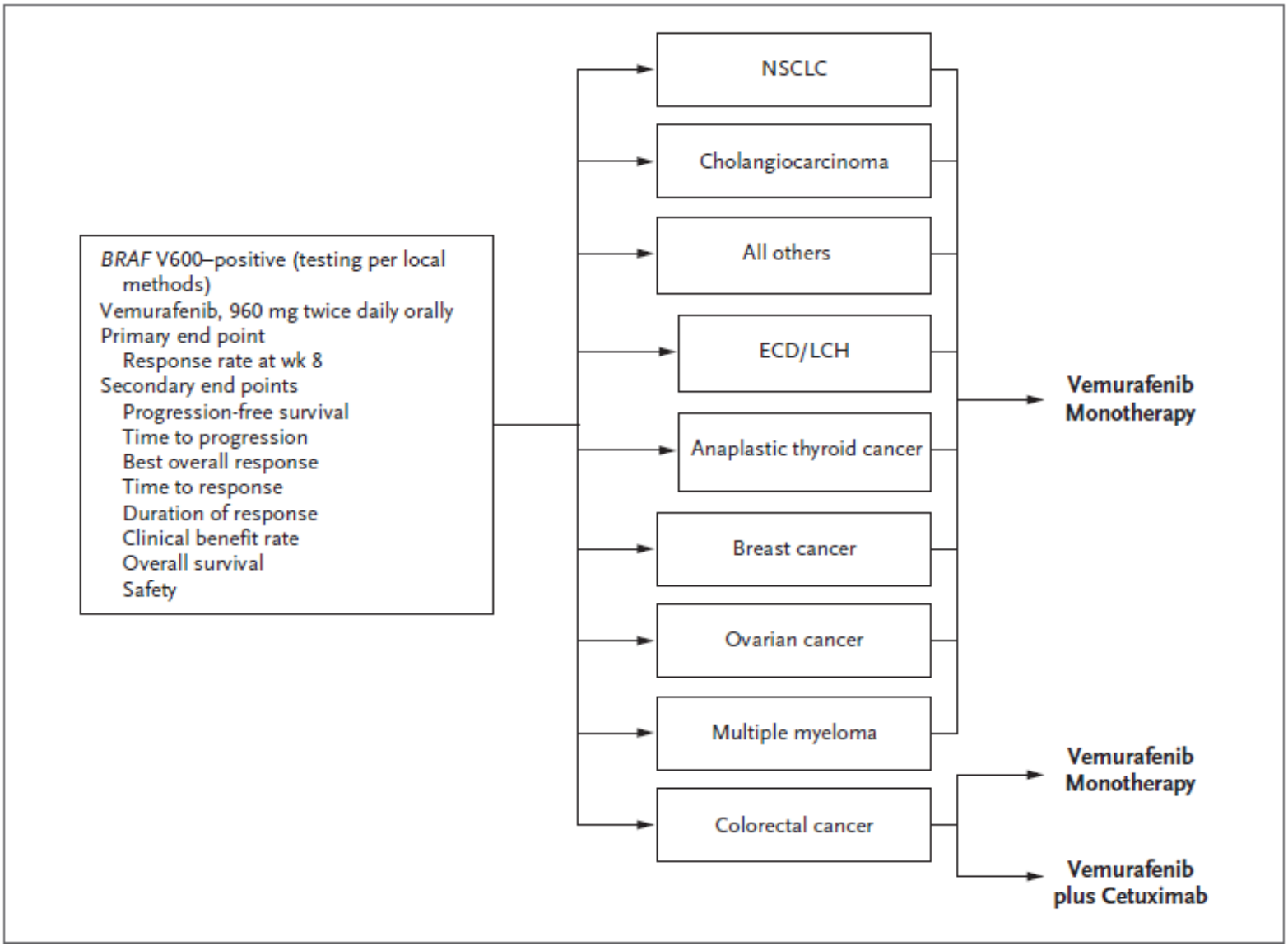
Setting: Common mutations present in multiple tumour types.

Aim: To find a **targeted therapy** for the common mutation

Solution Using biomarker(s) to screen patients and recruit those harbouring a common mutation



VE-BASKET Trial (Hyman et al. (2015))



Notation

- **Binary responses**: $Y_k \sim \text{Binomial}(n_k, p_k)$ total responses in basket k.
- n_k is the baskets sample size and p_k is the unknown response rate we are trying to estimate.
- q_0 = null response rate, q_1 = target response rate.
- Hypotheses: $H_0 : p_k \leq q_0$ vs. $H_1 : p_k > q_0$ for $k = 1, \ldots, K$.
- Treatment is deemed effective in basket k if $\mathbb{P}(p_k > q_0 | D) > \Delta_\alpha$
  - Δ_α is chosen based on type I error considerations.

Terminology

- **Type I Error** - Falsely conclude an ineffective treatment is effective.
  - $\mathbb{P}(p_k > q_0|D) > \Delta_\alpha$ when p_k is in fact null.
- **Power** - Correctly conclude an effective treatment is effective.
  - $\mathbb{P}(p_k > q_0|D) > \Delta_\alpha$ when p_k is in fact effective.
- **Homogeneous** - Baskets with identical or similar response rates.
- **Heterogeneous** - Baskets with differing response rates.

Bayesian Information Borrowing

Basket trials typically have **small sample sizes** within baskets $\Rightarrow$ **lack of statistical power**

With the **common genetic mutation** it may be assumed that patients will **respond similarly to the treatment**

Analysis strategies:

- Stratified analysis for each basket
- Complete pooling across all baskets
- Adaptive borrowing of information across baskets

Exchangeability

Information borrowing utilises the **exchangeability assumption**

The random variables $\theta_1, \theta_2, \ldots, \theta_K$ are exchangeable if

$$f_{\theta_1,\theta_2,\ldots,\theta_K}(t_1, t_2, \ldots, t_K) \overset{distr.}{=} f_{\theta_{\pi_1},\theta_{\pi_2},\ldots,\theta_{\pi_K}}(t_1, t_2, \ldots, t_K),$$

for any permutation $(\pi_1, \ldots, \pi_K)$ of the indices $\{1, 2, \ldots, K\}$.

It can be shown that

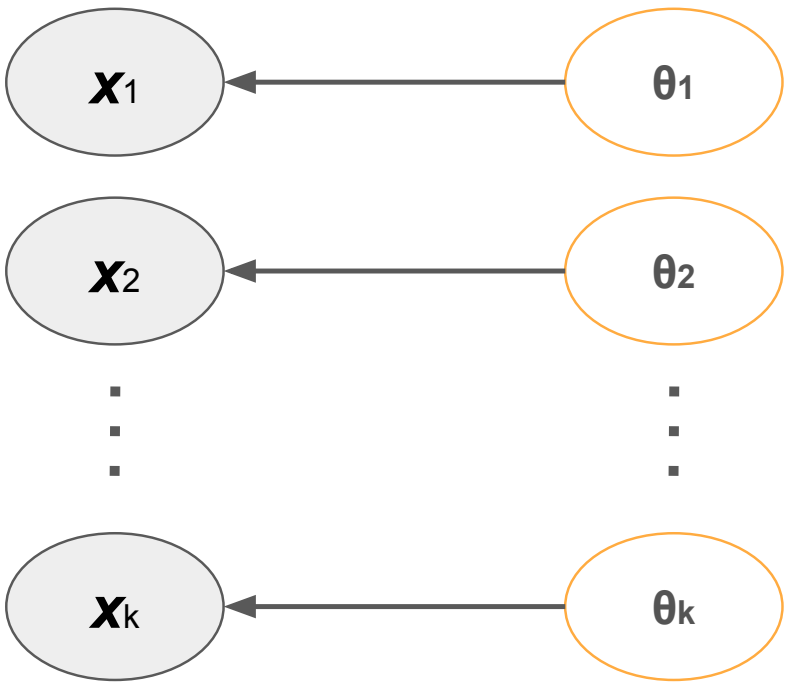
(1) i.i.d. $\Longrightarrow$ exchangeability,

(2) exchangeability $\Longrightarrow$ identically distributed.

Bayesian Hierarchical Model (BHM) (Berry et al. (2013))

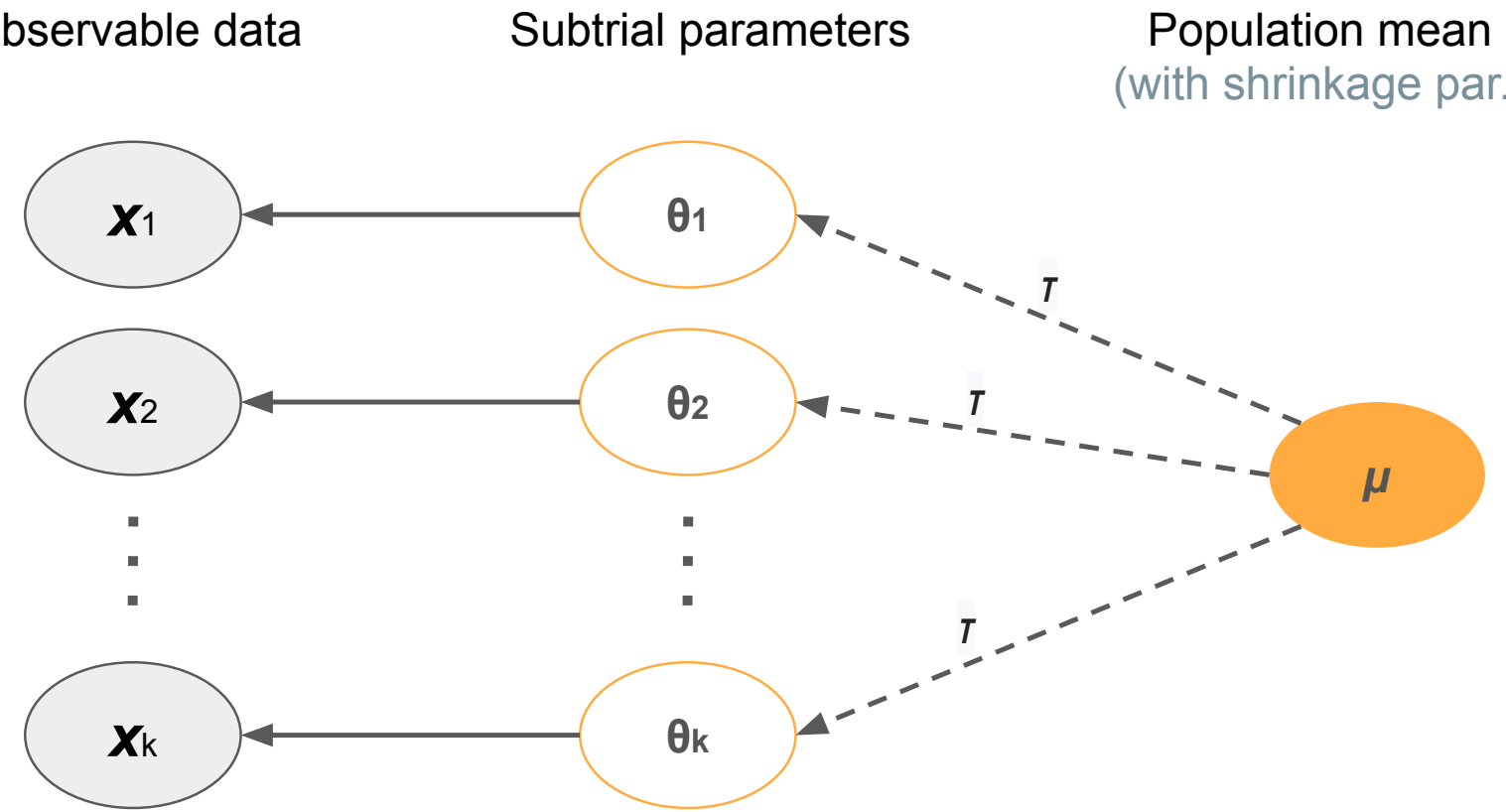
Let $\theta_k = \text{logit}(p_k)$

Observable data Subtrial parameters



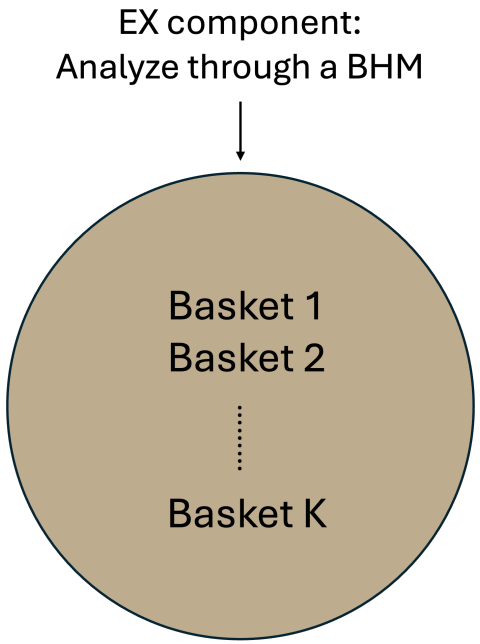
Bayesian Hierarchical Model (BHM) (Berry et al. (2013))

Let $\theta_k = \text{logit}(p_k)$



Bayesian Hierarchical Model (BHM) (Berry et al. (2013))

Assumes **exchangeability across all baskets** $\Rightarrow$ borrows information between all baskets.



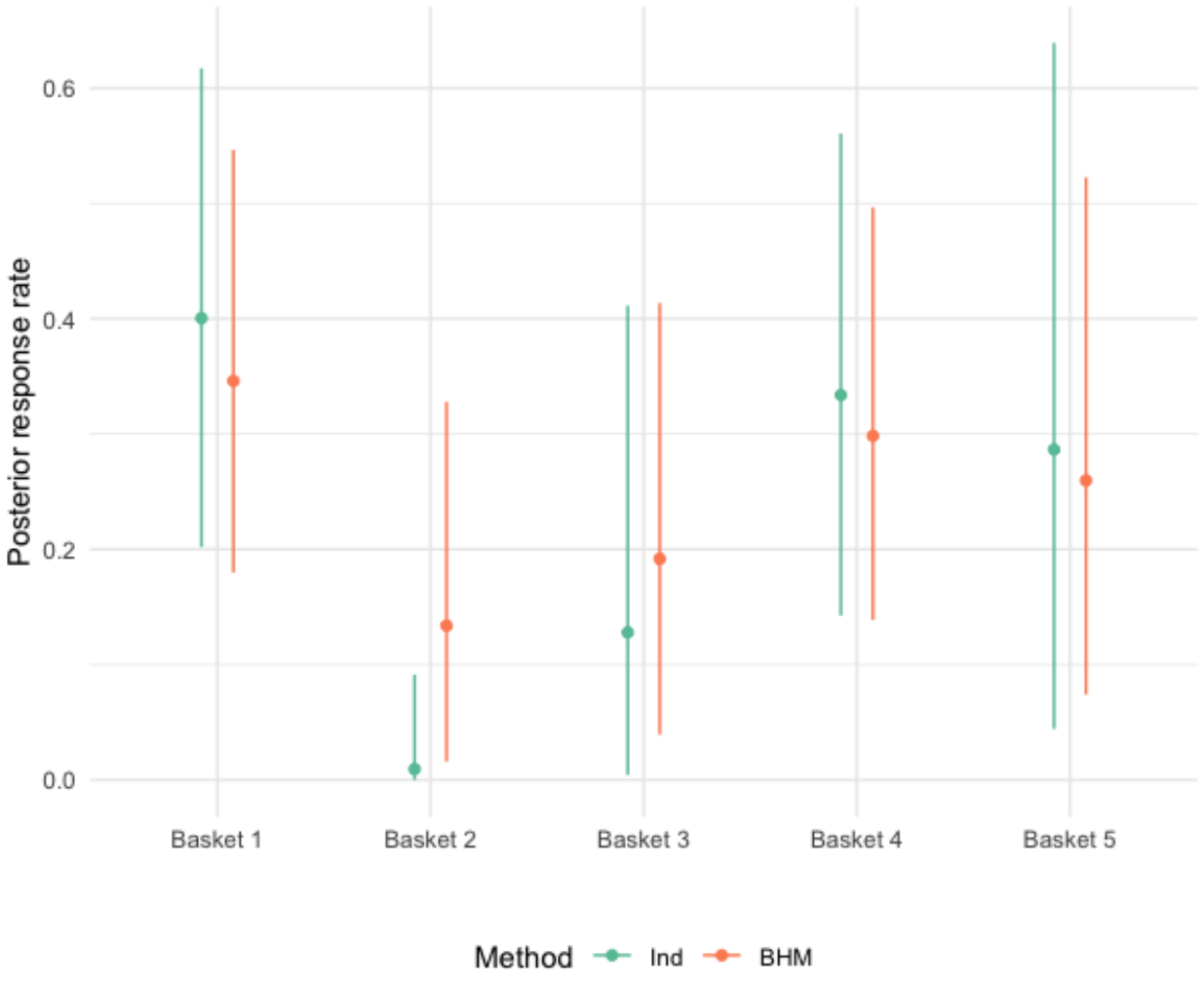
$$
\begin{aligned}
Y_k &\sim \text{Binomial}(n_k, p_k), \quad k = 1, \ldots, K \\
\theta_k &= \text{logit}(p_k) \sim N(\mu, \sigma^2), \\
\mu &\sim N(\text{logit}(q_0), \nu_\mu), \\
\sigma &\sim g(\cdot)
\end{aligned}
$$

Improves power but **inflates error rates** when one or more baskets have heterogeneous responses.

Example

Data

Basket	*y*	*n*
1	7	20
2	0	10
3	1	8
4	6	18
5	2	7



Calibrated BHM (CBHM) (Chu & Yuan (2018))

Rather than placing a prior $g(\cdot)$ on σ, define it as a **function of a measure of homogeneity across baskets:**

$$\sigma^2 = \exp\{a + b \log(T)\},$$

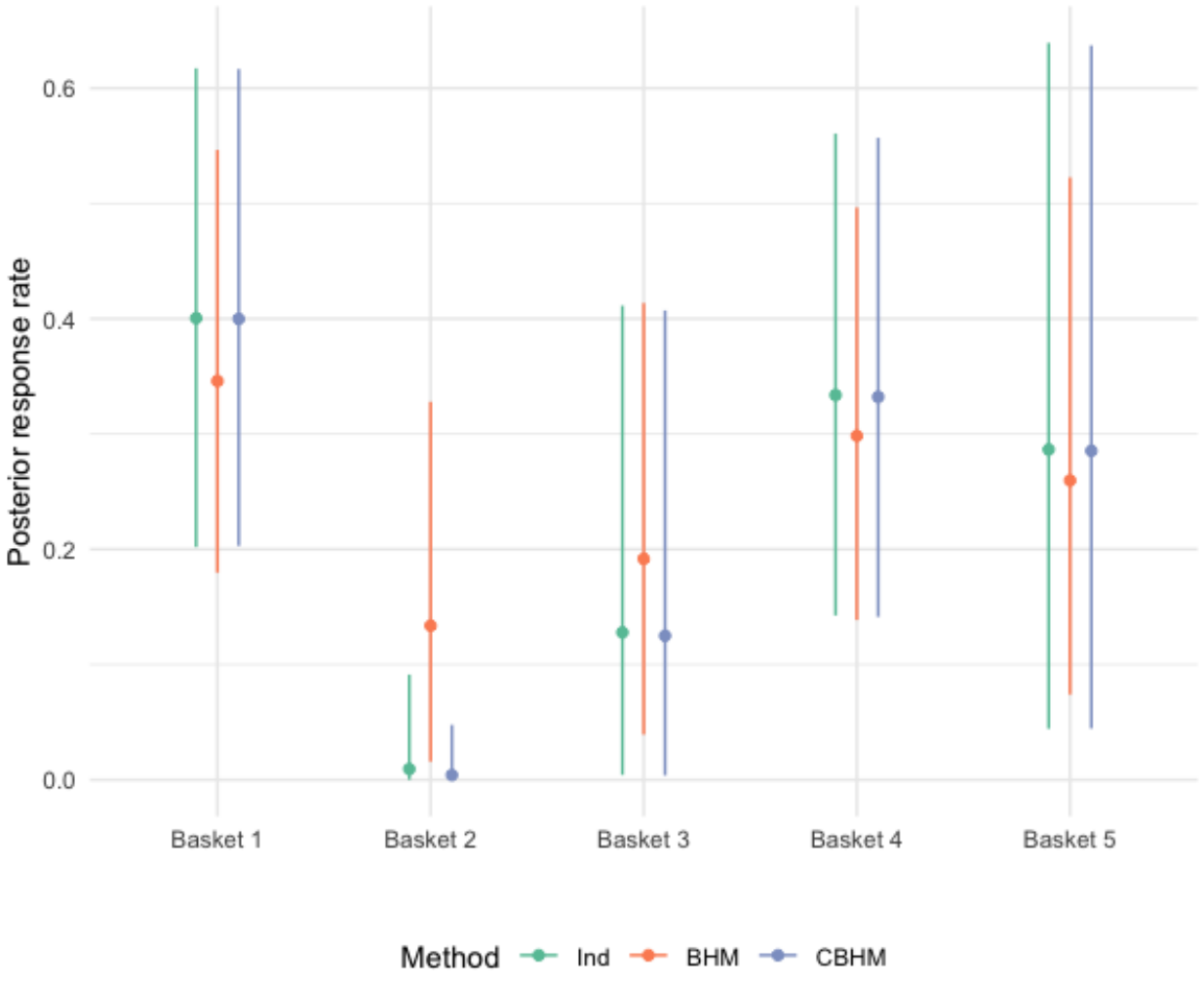
- T: Chi-squared test statistics for homogeneity;
- a, b: Tuning parameters.

Takes '**strong' definition of heterogeneity**: If at least one basket has a heterogeneous response, all are deemed heterogeneous and no borrowing occurs.

Example

Data

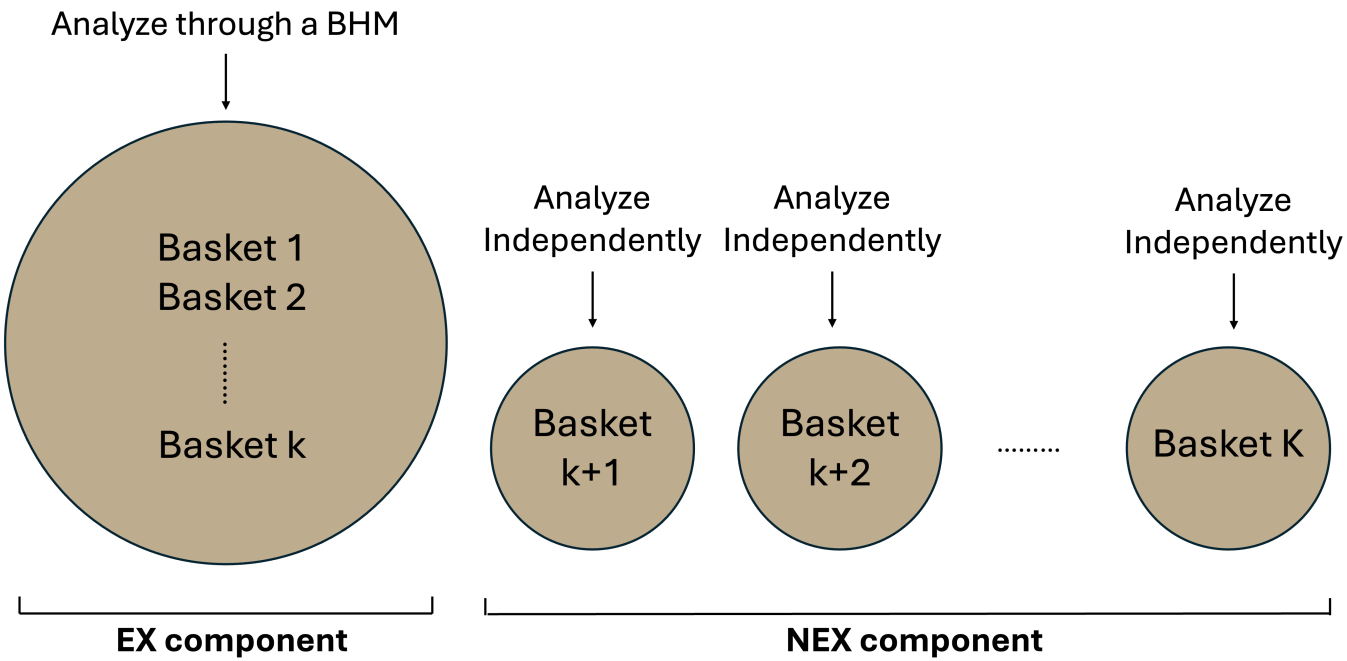
Basket	*y*	*n*
1	7	20
2	0	10
3	1	8
4	6	18
5	2	7



Exchangeability-Nonexchangeability (EXNEX) Model (Neuenschwander et al. (2016))

Relaxes exchangeability assumption by having two components:

- **Exchangeable** (EX) - borrow between all baskets assigned. Baskets are assigned with prior probability π_k.
- **Nonexchangeable** (NEX) - independent analysis on all baskets assigned. Baskets are assigned with prior probability $1 - \pi_k$.



Exchangeability-Nonexchangeability (EXNEX) Model (Neuenschwander et al. (2016))

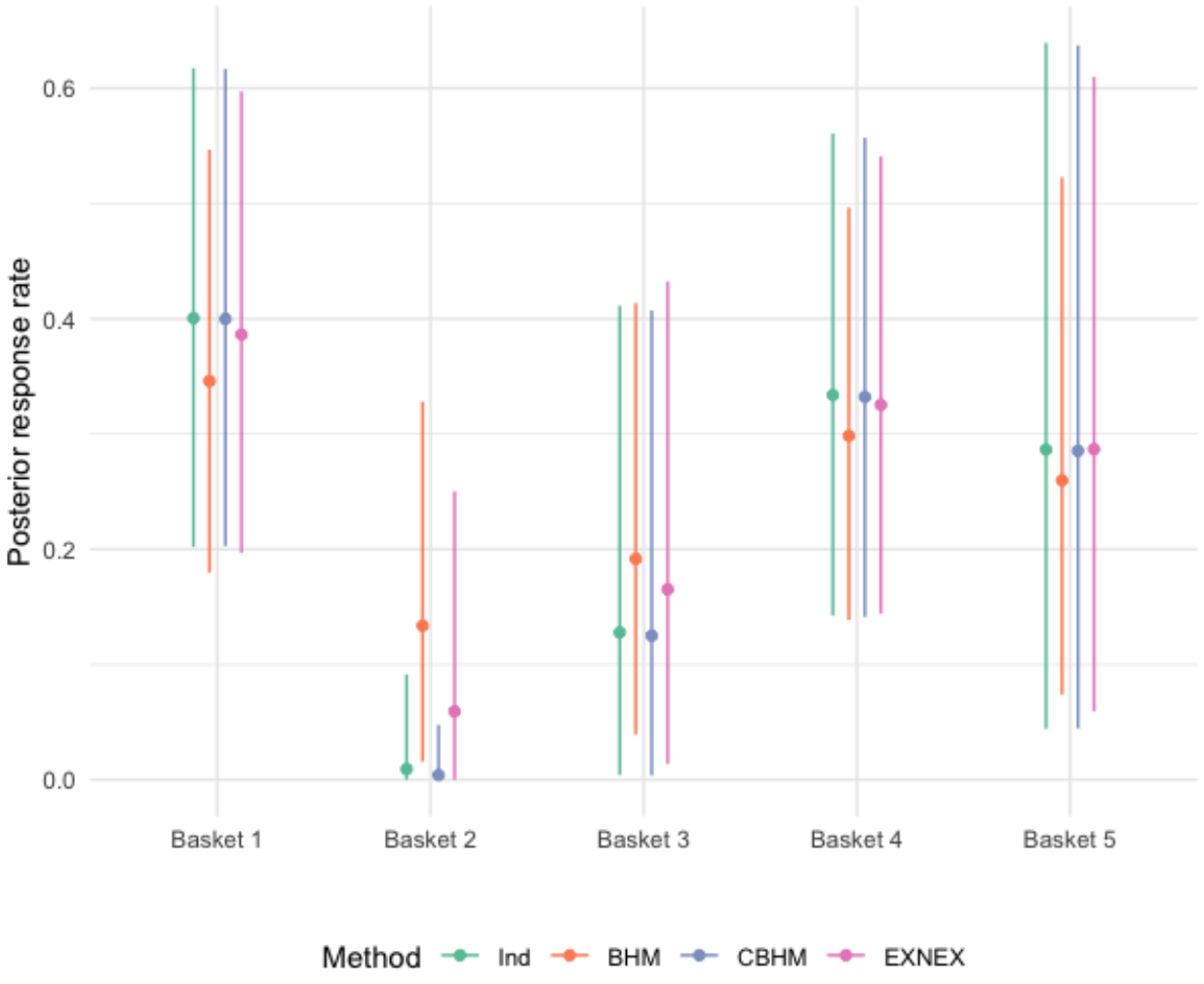
$$
\begin{aligned}
Y_k &\sim \text{Binomial}(n_k, p_k), \\
\theta_k &= \log\left(\frac{p_k}{1-p_k}\right), \\
\theta_k &= \delta_k M_{1k} + (1-\delta_k) M_{2k}, \\
\delta_k &\sim \text{Bernoulli}(\pi_k),
\end{aligned}
\qquad
\begin{aligned}
M_{1k} &\sim \text{N}(\mu, \sigma^2), && \text{(EX)} \\
\mu &\sim \text{N}(\text{logit}(q_0), \nu_\mu), \\
\sigma &\sim g(\cdot), \\
M_{2k} &\sim N(m_k, \nu_k). && \text{(NEX)}
\end{aligned}
$$

- EX is a **Bayesian hierarchical model**.
- NEX is a slightly informative basket specific prior.
- π_k typically set to 0.5 a priori for all k.
- Resembles the **rMAP** (Bayesian dynamic borrowing).

Example

Data

Basket	*y*	*n*
1	7	20
2	0	10
3	1	8
4	6	18
5	2	7



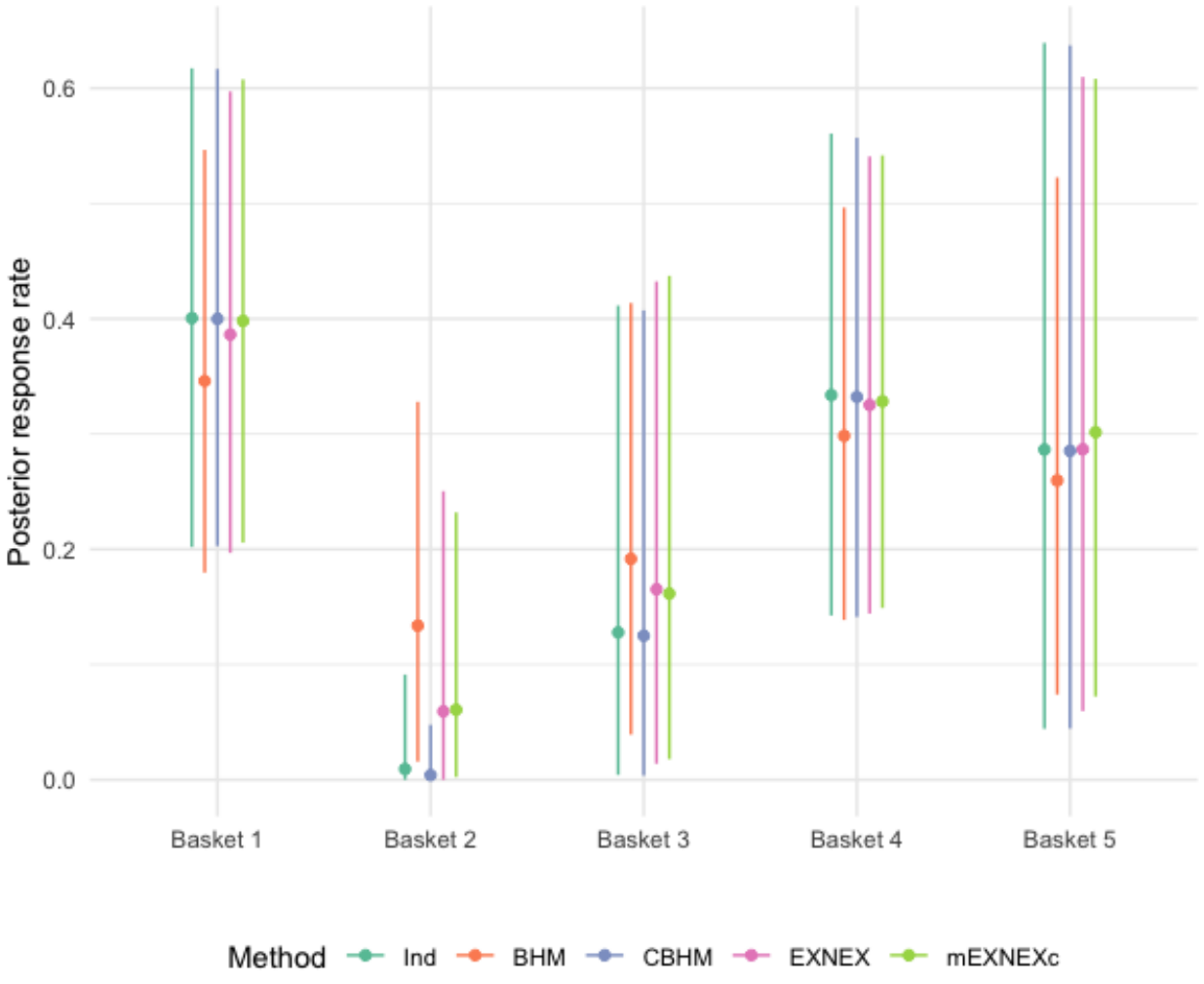
Modified EXNEX (mEXNEX_c) (Daniells et al. (2023))

- Use a **data-driven** approach to set π_k.
- Encourage more borrowing between baskets whose response data is homogenous and less borrowing when heterogeneous.
- Two-step procedure:
  1. If $\min_{j \neq k}\{|\hat{p}_k - \hat{p}_j|\} > c \;\Rightarrow\; \pi_k = 0$, i.e. analyze independently. Let S = Set of all baskets not excluded for heterogeneity.
  2. $\pi_k = \sum_j \frac{1-h_{kj}}{|S|-1}, \;\; k \neq j, \; k, j \in S$, i.e. borrowing probability is the mean pair-wise **Hellinger distance** between non-heterogeneous baskets.

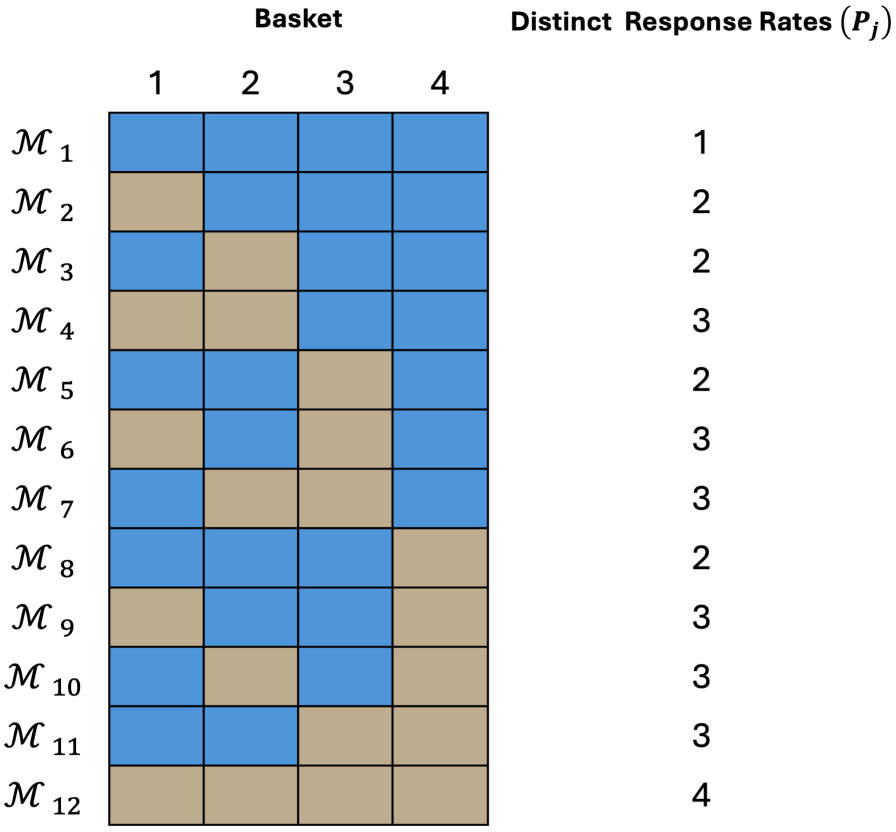
Example

Data

Basket	*y*	*n*
1	7	20
2	0	10
3	1	8
4	6	18
5	2	7



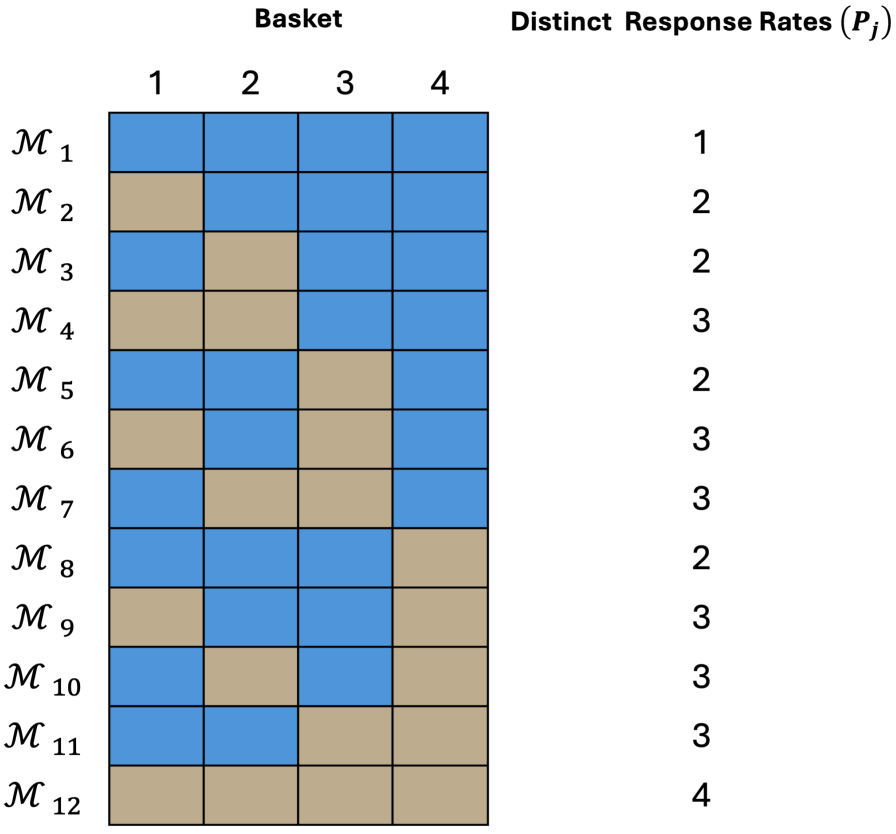
Bayesian Model Averaging (BMA) (Psioda et al. (2021))



4 basket example and all possible models. Blue represents the baskets assignment to the EX component, brown to the NEX.

- Baskets are assigned to the EX or the NEX component.
- Find all permutations for basket assignment to the EX and NEX components.
- Each permutation is a model, denoted $\mathcal{M}_j$.

Bayesian Model Averaging (BMA) (Psioda et al. (2021))



4 basket example and all possible models. Blue represents the baskets assignment to the EX component, brown to the NEX.

- Within the EX component, responses are pooled. There is one response rate for the EX component.
- Each basket in the NEX component has a distinct response rate.
- P_j is the total number of distinct responses in model j.

Bayesian Model Averaging (BMA) (Psioda et al. (2021))

Rather than basing inference on a single model, we take the **average** across several plausible models.

- Prior on each model, $f(\mathcal{M}_j)$, e.g. $f(\mathcal{M}_j) \propto P_j^2$.
- Prior on the response rates given a model j, $f(p_k|\mathcal{M}_j)$, e.g. $f(p_k|\mathcal{M}_j) \sim \text{Beta}(a_0, b_0)$.
- Compute posterior the $f(p_k|\mathcal{M}_j, D)$ given response data D.
- Compute the posterior $f(\mathcal{M}_j|D)$.

$$\mathbb{P}(p_k > x)|D) = \sum_j \mathbb{P}(p_k > x|\mathcal{M}_j, D) f(\mathcal{M}_j|D)$$

Simulation Study

- Total of $K = 5$ baskets with $n_k = 13$ patients in each.
- $q_0 = 0.15$ and $q_1 = 0.35$.

	p_1	p_2	p_3	p_4	p_5
Scenario 1	0.15	0.15	0.15	0.15	0.15
Scenario 2	0.45	0.15	0.15	0.15	0.15
Scenario 3	0.45	0.45	0.45	0.45	0.15
Scenario 4	0.45	0.45	0.45	0.45	0.45

- Each method is **calibrated** to have 10% type I error rate under the global null scenario (i.e. scenario 1).

What do we Mean by Calibration?

1. Sample response data, Y_k from Binomial(n_k, p_k), where p_k is null in all K baskets.
2. Fit a model to the response data.
3. Compute the posterior probability $\mathbb{P}(p_k > q_0|D)$.
4. Repeat steps 1-3 N times, storing each posterior probability.
5. Δ_α is set as the 90% quantile of the posterior probabilities.

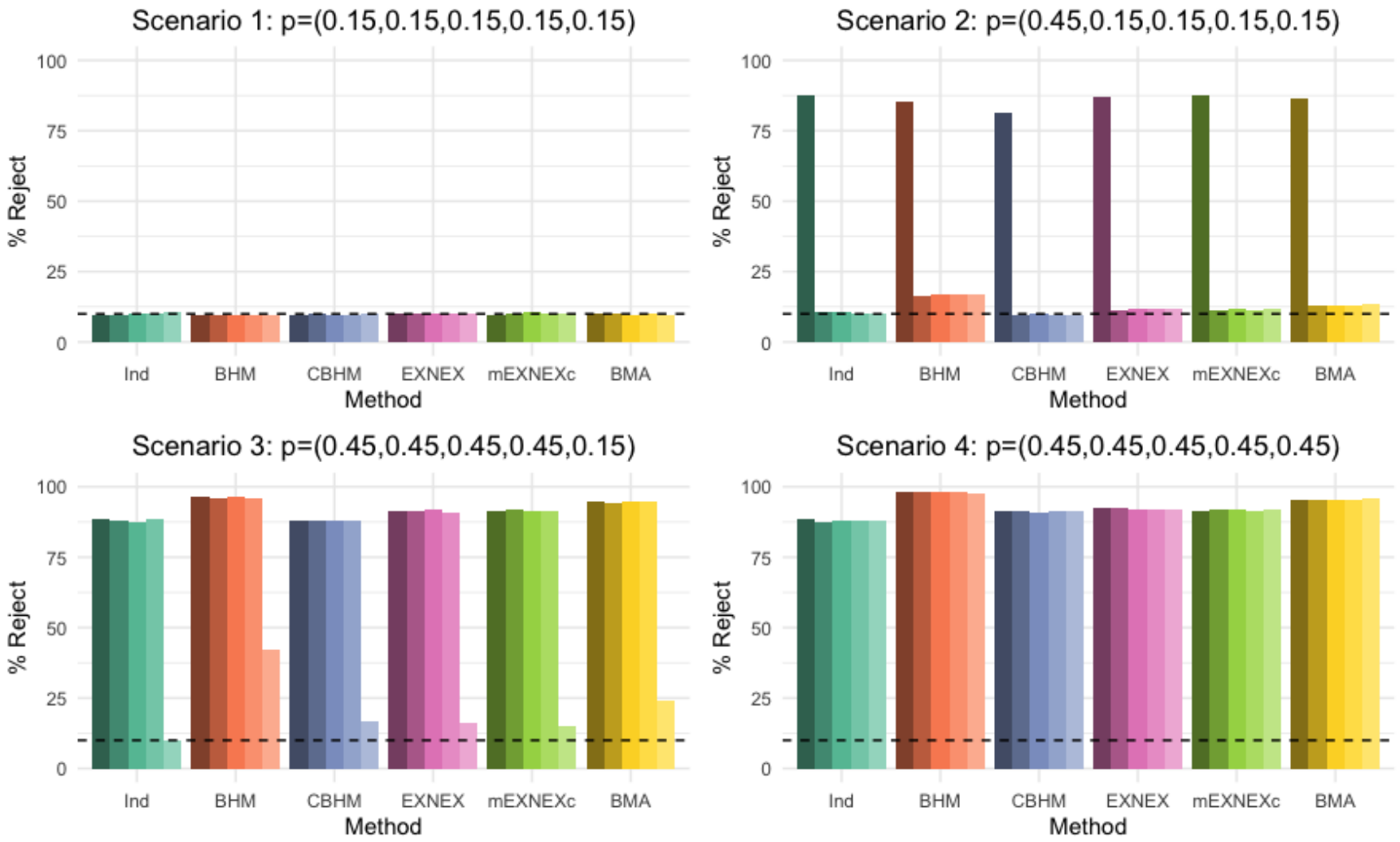
Ensures the treatment is deemed effective in 10% of simulation runs under the null scenario $\Rightarrow$ 10% type I error rate control.

Alternative: **Robust Calibration Procedure** (RCaP, Daniells et al. (2025)) controls type I error across numerous scenarios.

Simulation Study

- Metric of interest: % times the null hypothesis is rejected.
  - When the null is true this is the **Type I error rate**.
  - When the null is false this is the **Power**.

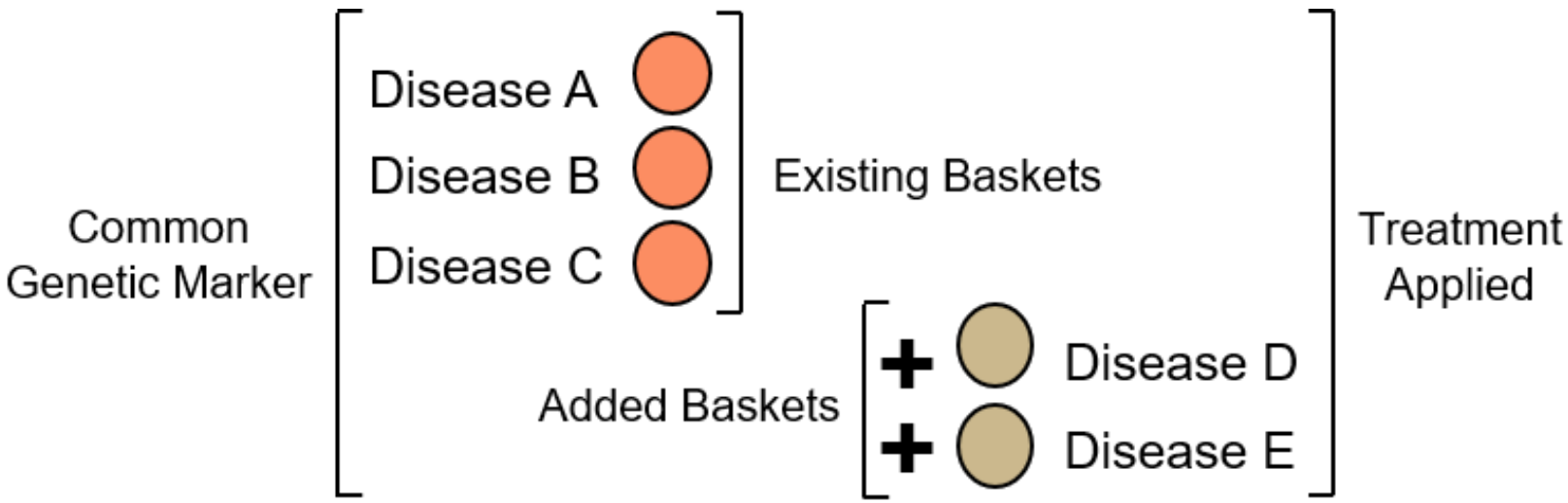
Simulation Results



Adding Baskets to an Ongoing Trial Daniells et al. (2025)

Adaptive design features can be incorporated into basket trials.

E.g. Interim analyses, sample size adjustment, stopping for futility or **adding baskets** to an ongoing trial.



Adding Baskets to an Ongoing Trial

Approaches:

1. **IND** - Analyse the new basket as independent of existing baskets. Borrow information between just existing baskets.
2. **UNPL** - An unplanned addition. Don't acount for the new basket in calibration. For analysis, borrow between all baskets.
3. **PL1** - A planned addition. Borrow between all existing and new baskets.

The 'best' approach depends on the degree of agreement between new and existing baskets. Significant power can be gained via information borrowing between all baskets.

Discussion

- Strong justification (such as a common genetic make-up or disease trait) $\Rightarrow$ borrowing
- In data analysis, borrowing leads to higher statistical power than no borrowing.
  - Can come with a risk of type I error rate inflation under most borrowing methods.
  - Error inflation can be limited through calibration of efficacy criteria.
- By taking into account borrowing in the design stage, the sample size could be reduced to achieve the same power as a design without borrowing (Zheng et al. (2023)).

References (1)

- Ferrarotto R, Redman M, Gandara D, Herbst R, Papadimitrakopoulou V. (2015) Lung-MAP-framework, overview and design principles. Chinese Clinical Oncology. 4(3), 36.
- Hyman D, Puzanov I, Subbiah V, et al. (2017) Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. New England Journal of Medicine. 373(8):726-736.
- Berry S, Broglio K, Groshen S, Berry D. (2013) Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology trials. Clin Trials. 10(5):720-734.
- Chu Y, Yuan Y. (2018) A Bayesian basket trial design using a calibrated Bayesian hierarchical mode. Clin Trials. 15(2):149-158
- Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. (2016) Robust exchangeability designs for early phase clinical trials with multiple strata. Pharmaceutical Statistics 15(2):123-34.

References (2)


Psioda MA, Xu J, Jiang Q, Ke C, Yang Z, Ibrahim JG. (2021). Bayesian adaptive basket trial design using model averaging. Biostatistics. 22(1):19-34.


Daniells L, Mozgunov P, Barnett H, Bedding A, Jaki T. (2023) A comparison of Bayesian information borrowing methods in basket trials and a novel proposal of modified exchangeability-nonexchangeability method. Statistics in Medicine. 42(24):4392-4417.


Zheng H, Grayling M J, Mozgunov P, Jaki T, Wason J M. (2023) Bayesian sample size determination in basket trials borrowing information between subsets. Biostatistics. 24(4):1000-1016.


Daniells L, Mozgunov P, Barnett H, Bedding A, Jaki T. (2025) How to add baskets to an ongoing basket trial with information borrowing. Statistical Methods in Medical Research.