MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Basket trials – a powerful tool for precision medicine trials

Thomas Jaki & Pavel Mozgunov

Bayes Pharma 2023

October 25, 2023

Outline

09:00 – 09:45	Introduction to Basket Trials and Bayesian Inference	Jaki
09:45 – 10:30	Bayesian Borrowing Models	Mozgunov
10:30 – 10:50	Coffee Break	
10:50 – 12:00	Practical	Mozgunov
12:00 – 12:30	How to design a basket trial	Jaki

Introduction to Basket Trials

Master protocols

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

- Platform trials
- Basket trials
- Umbrella trials

Changing paradigms in cancer research

Categorisation of cancer

Anatomic location of the primary tumor

Breast cancer
Colorectal cancer
Lung cancer, etc.

Molecularly defined subtypes

'Cancer is caused by the alterations in normal genes.'

Anti-cancer treatment

Cytotoxic chemotherapy

'Toxic to cells': disrupt the way cancer cells grow and divide
Can often affect healthy cells

Cytostatic drugs

Targeted therapy
Immunotherapy

Groundbreaking approval of immunotherapy

- Keytruda (pembrolizumab), an antibody that attaches to a molecule called PD-1
- To treat unresectable or metastatic solid tumors with a specific biomarker
- First FDA approval of anti-cancer treatment based on biomarkers rather than tumour location
- Read more: `www.fda.gov/newsevents/newsroom/pressannouncements/ucm560167.htm`

Biomarker-driven designs

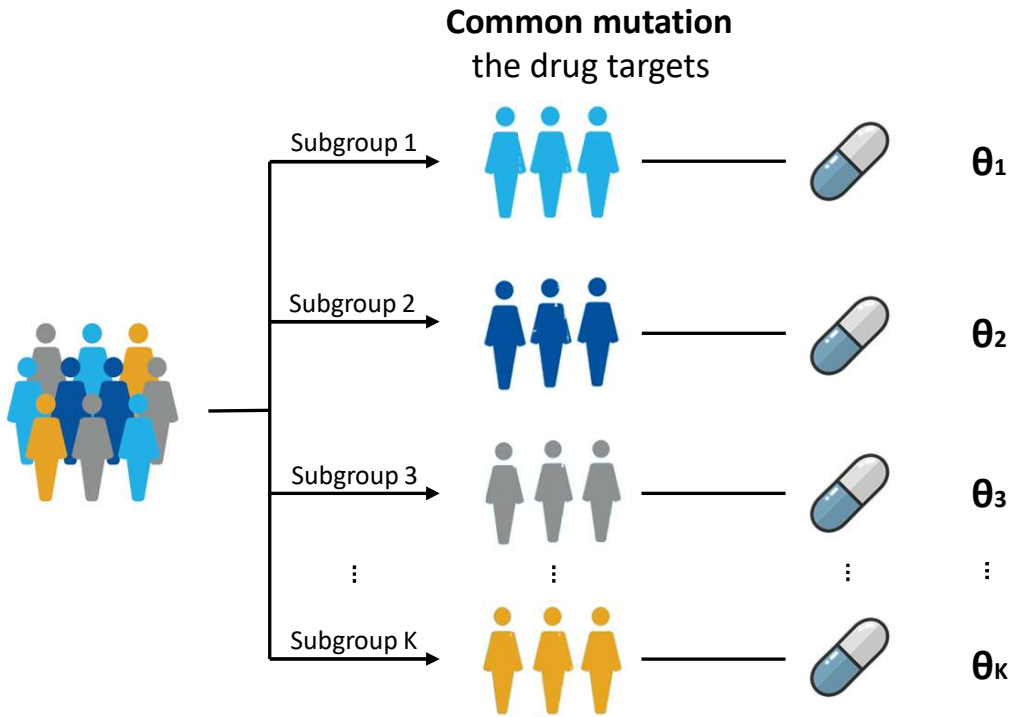
- Molecular profiling at the individual patient level became feasible and affordable
- **Biomarker**: e.g., measurable indicator of biological properties or genetic aberration
- More trials are now biomarker-driven
- Choice of the design and analysis relies on the biomarker's nature, e.g., prognostic or predictive
- Well-known types: (adaptive) enrichment designs, master protocols
- **Key objective**: increased efficiency for drug development when target-drug links exist

Basket trials

Setting: Common characteristic (e.g. mutation) present in multiple tumour types.

Aim: To develop **targeted therapies**

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

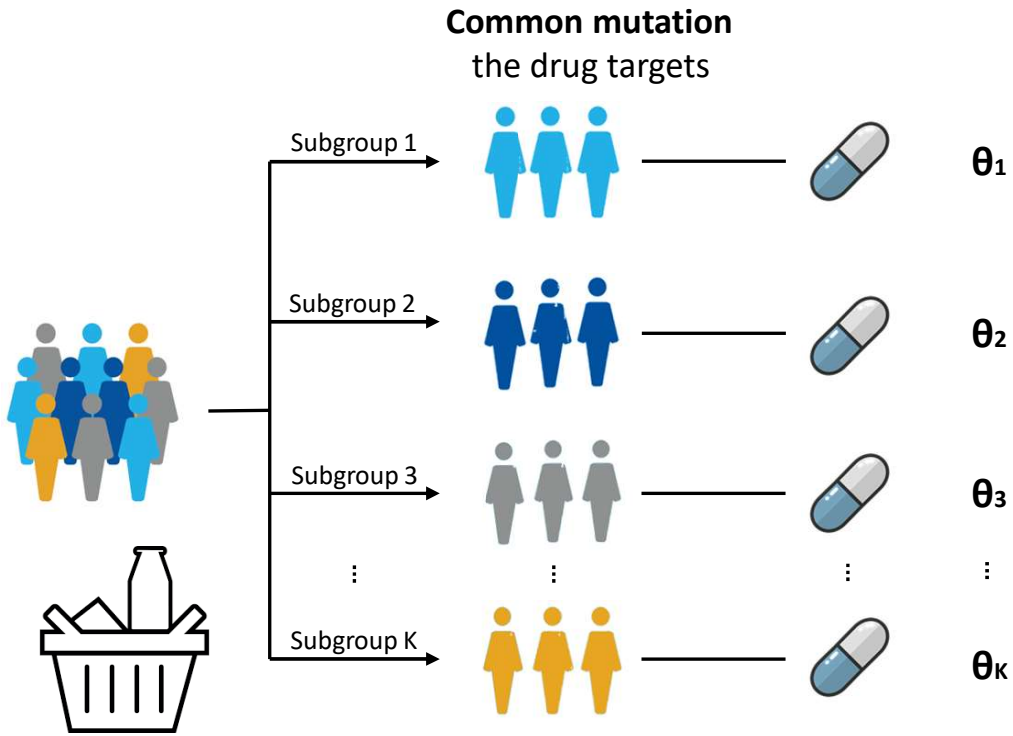


Basket trials

Setting: Common characteristic (e.g. mutation) present in multiple tumour types.

Aim: To develop **targeted therapies**

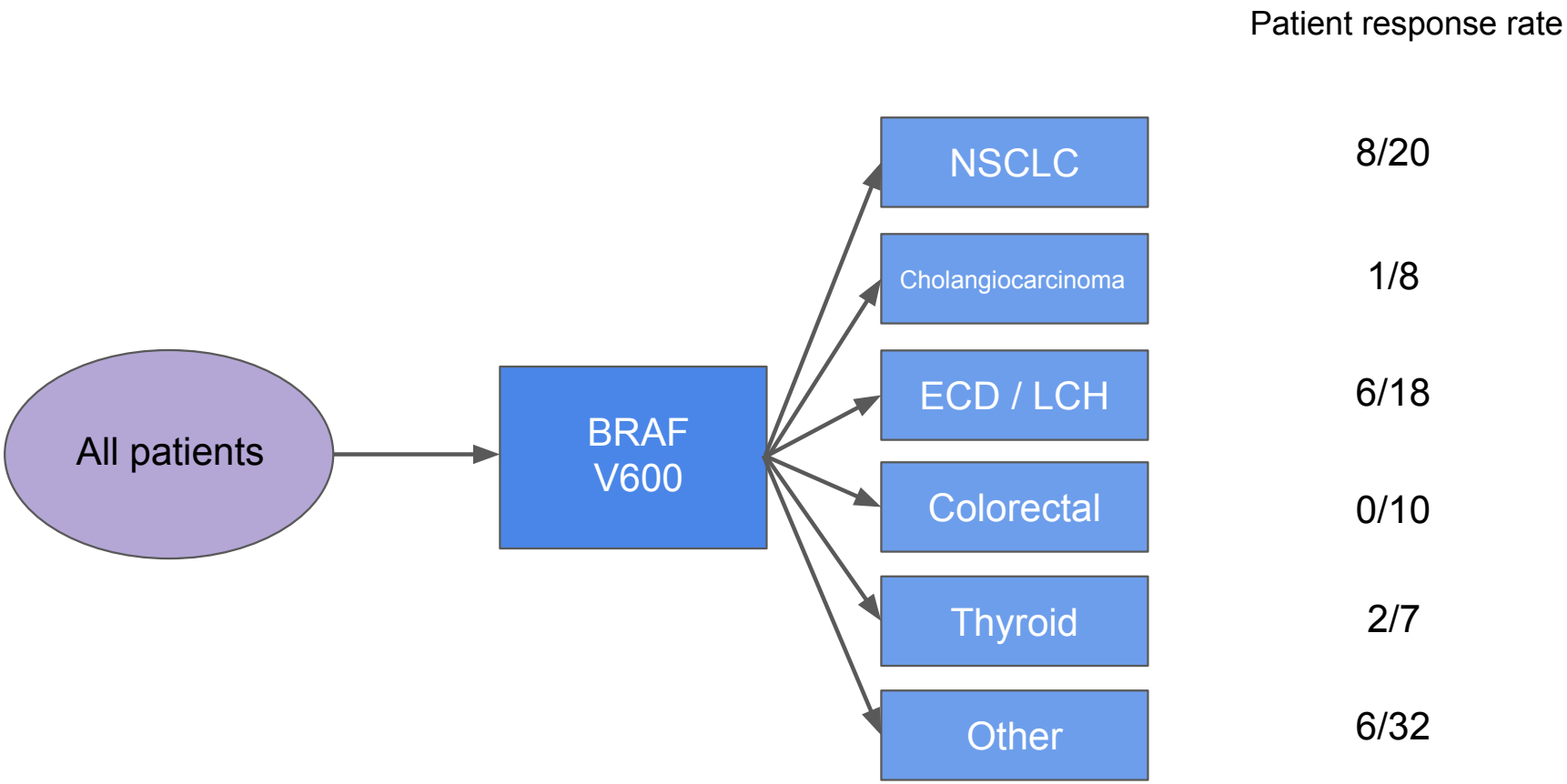
Solution Using biomarker(s) to screen patients and recruit those harbouring a common characteristic (mutation)



Basket trials in oncology – An example

Hyman *et al.* (2015) reported a basket trial, which has been designed to evaluate the efficacy of vemurafenib in BRAF-V600.

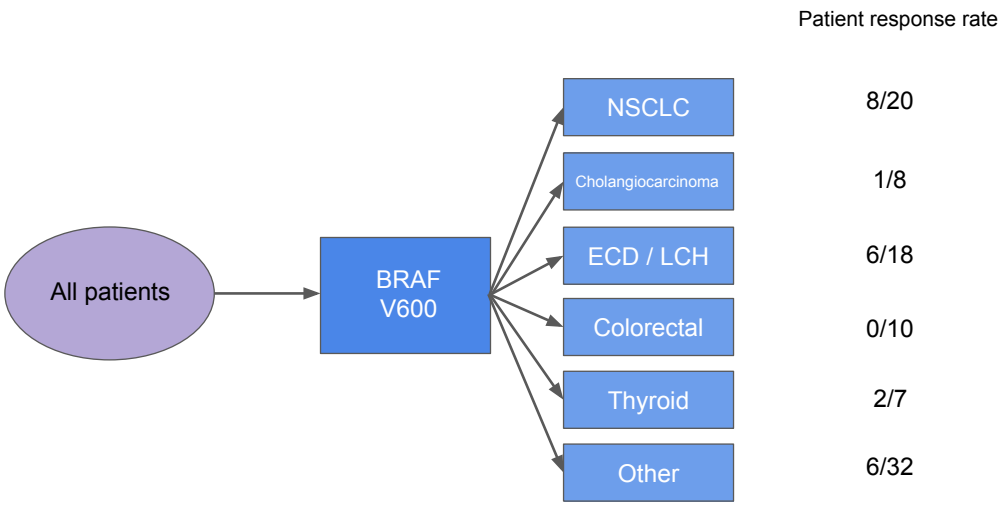
A total of 122 patients with BRAF-V600 mutations were enrolled, of which 95 entered the 6 modules.



Borrowing of information between modules

With the **common genomic mutation** targeted by the investigational drug, one may expect that

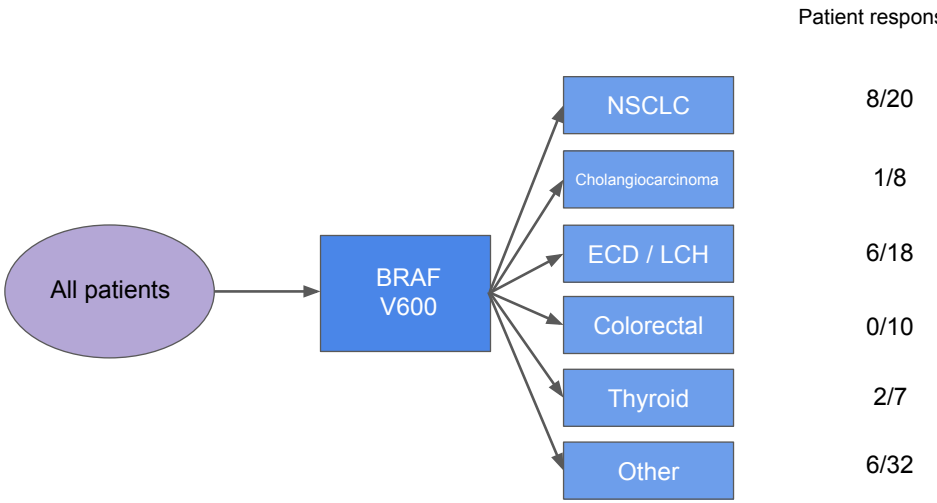
... some patient subgroups will **respond similarly**.



Borrowing of information between modules

With the **common genomic mutation** targeted by the investigational drug, one may expect that

... some patient subgroups will **respond similarly**.



Potential analysis strategies:

- Stand-alone analyses
- Complete pooling
- Borrowing of information

Basket trials outside of oncology

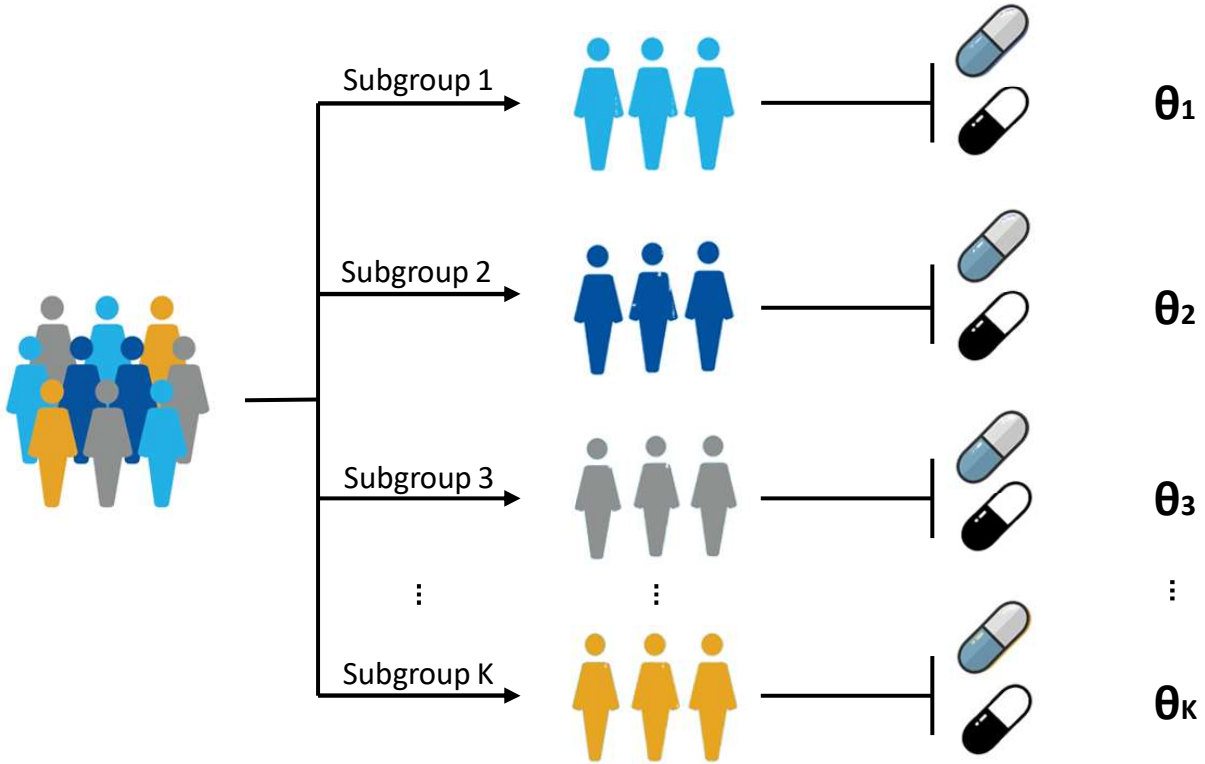
Basket trials evaluate a treatment for multiple disease indications with **a common characteristic**, e.g.

- a genomic biomarker;
- mechanism of drug activity;
- clinical symptom that the treatment targets

There is a **great need** in more efficient study designs utilising this feature of a common characteristic outside of oncology.

- Neurodegenerative (NDD) and neuromascular diseases: drugs that address aspects of biology or symptoms shared with other NDDs (Cummings et al. 2022)
- Centronuclear myopathies (CNMs): The disorders share a set of common pathologiesy and phenotypes (Fourage et al. 2021)
- Rare metabolic disorders with causes within the same pathway;
- and many more...

Randomised controlled basket trials



Introduction to Bayesian Inference

Bayesian inference

Why being Bayesian for Basket trials?

Bayesian inference provides a formal approach for incorporating information from additional sources, e.g.

- expert knowledge;
- historical data;
- **other baskets in the trial**;

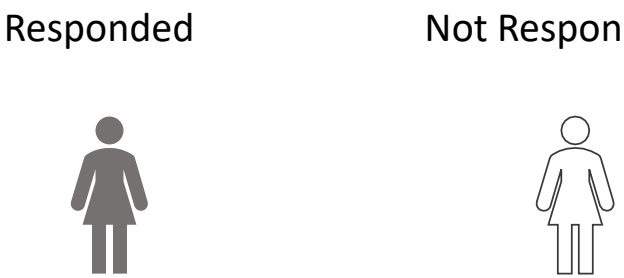
Concept of Bayesian inference:

- Our (prior) knowledge about a common characteristic of baskets is incorporated through an *appropriately chosen* prior.

Example

- Consider a **Phase II study**
- Question: does a new drug provide any benefits in terms of Response Rate (θ_i) in a given disease population (basket) *i*;

- **Binary Outcome**: Response vs No Response;



A Bayesian Approach

- Let θ_i be the parameter of interest, the probability of response (or a function of it) in basket i, that is a **random variable** in itself
- θ_i has a prior distribution $\pi(\theta_i)$ reflecting our uncertainty / knowledge about it
- Information about nature of θ_i comes from the sample $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ that has marginal density function $f(\mathbf{y})$
- The likelihood function $f(\mathbf{y}|\theta_i)$ is the distribution of $\mathbf{y}$ conditional on specific values of θ_i.
- Bayes' theorem:

$$\pi(\theta_i|\mathbf{y}) = \frac{f(\mathbf{y}|\theta_i)\pi(\theta_i)}{f(\mathbf{y})}$$

Stand-alone (Independent) Model

Independent/Stand-alone analysis is an approach that does not borrow information between baskets and instead conducts stratified analysis for each.

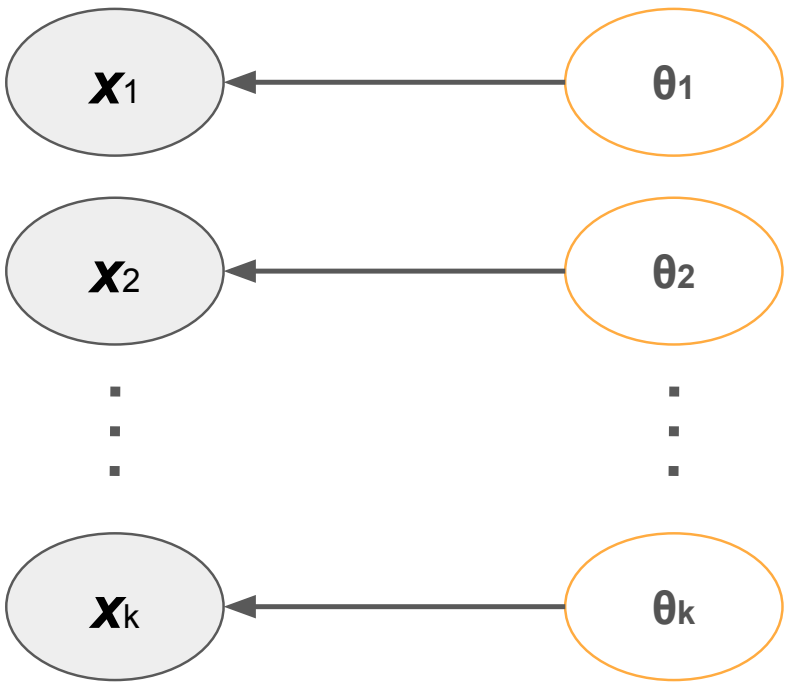
For each basket i,

$$\begin{aligned} Y_i &\sim \text{Binomial}(n_i, p_i), \\ \theta_i &= \log\left(\frac{p_i}{1-p_i}\right), \\ \theta_i &\sim \text{N}(m_i, \nu_i), \end{aligned} \tag{1}$$

where m_i and ν_i are parameters of the prior distribution (called **hyperparameters** to distinguish them from the parameters of the sampling space) for basket i.

Stand-alone (Independent) Model

Observable data Subtrial parameters

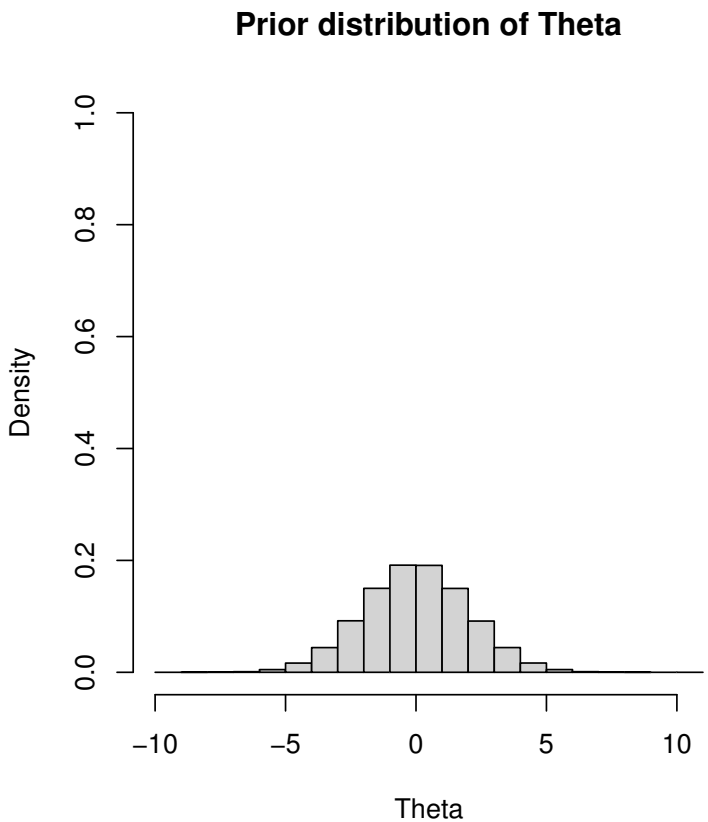


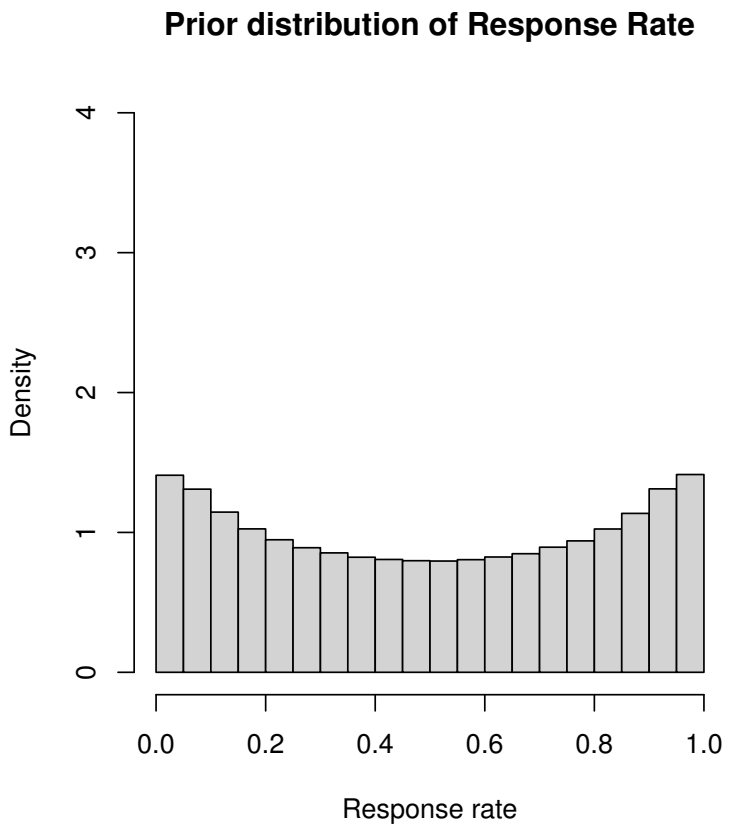
Prior distribution

Assume a weakly informative prior on parameter θ in Model (1)

$$\theta \sim \mathrm{N}(0.0, 4)$$

Samples from the prior distribution in Model (1) can be obtained





Likelihood function

Critical to Bayesian inference are likelihood functions. These

- serve to link the **sampling** space to the **parameter** space.

The probability:

$$f(\mathbf{y}|\theta_i) = L(\theta_i) = \prod_{l=1}^{n} f(y_l|\theta_i)$$

is a general expression for a likelihood function given *iid* data.

Example: Bernoulli likelihood

- The observation of 8 responses from 20 patients is a series of Bernoulli trials with $y = \{0, 1\}$. For any single trial l the probability mass function is

$$f(y_l|\,\theta) = \theta^{y_l}(1-\theta)^{1-y_l}$$

- The likelihood function is therefore:

$$L(\theta_i) = f(\mathbf{y}|\,\theta_i) = \prod_{l=1}^{n} \theta^{y_l}(1-\theta)^{1-y_l} = \theta^{\sum_l y_l}(1-\theta)^{n-\sum_l y_l}$$

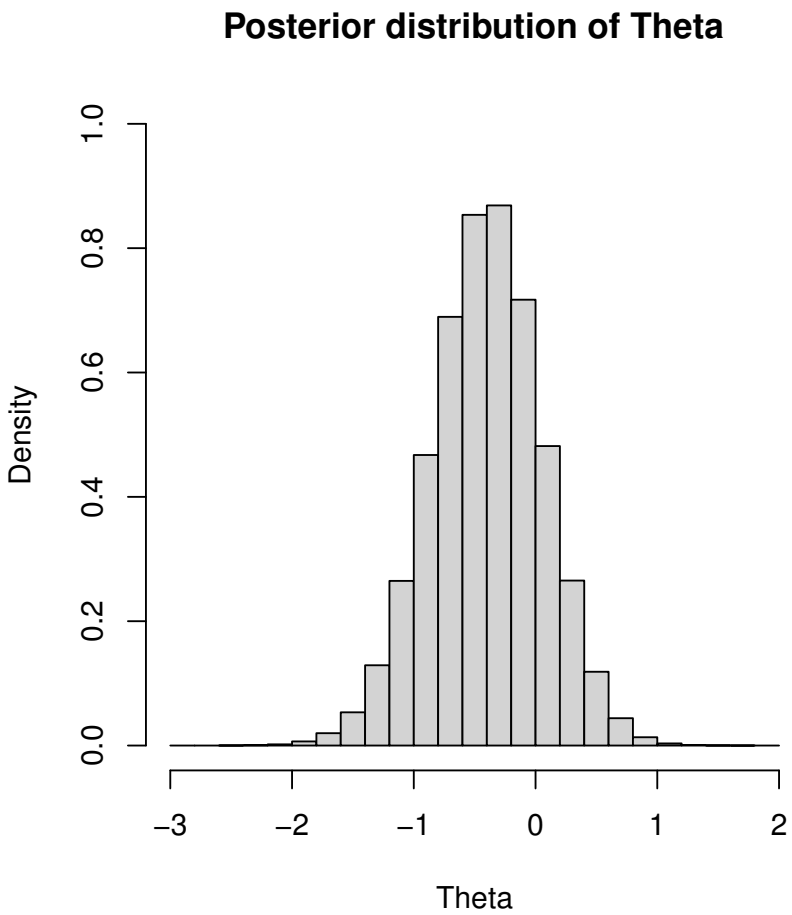
This is the functional form of the *Beta* distribution.

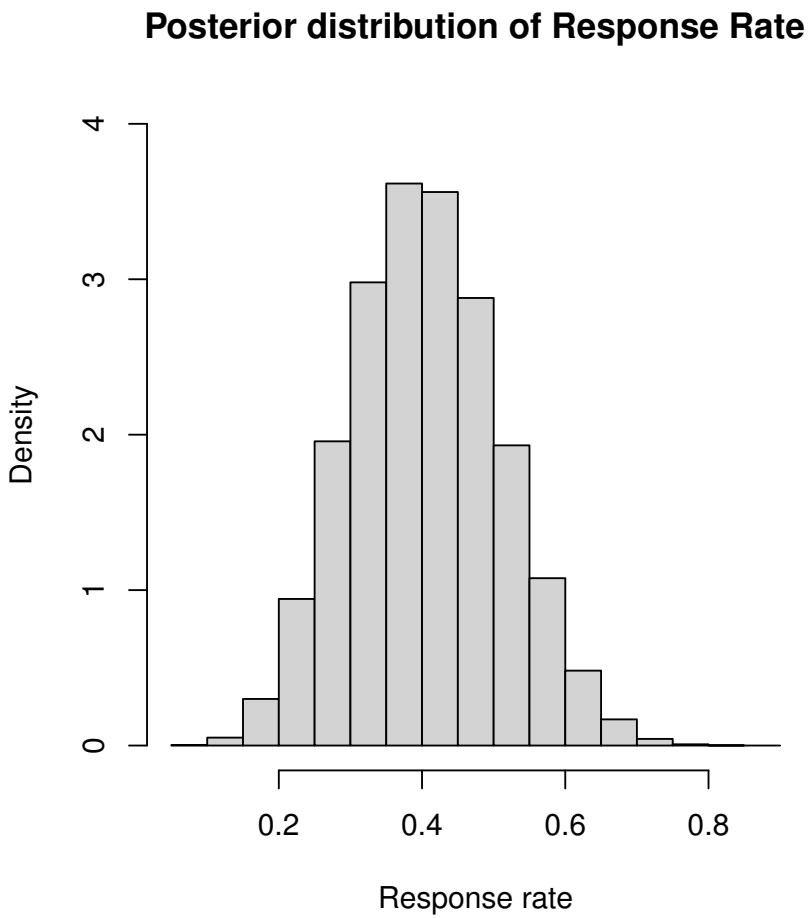
Note on conjugate prior and MCMC

- For the Bernoulli likelihood, if one chooses a prior in the same function form (i.e. the Beta prior distribution), then the posterior will be a Beta distribution again;
- This is called a **conjugate prior**;
- We have chosen a normal (on logit scale) prior to draw a parallel with other (more complicated) models;
- If the prior is not conjugate (as in our case), one can use the **Markov Chain Monte Carlo** (MCMC) to obtain the (approximation) of the posterior distribution of interest;
- The samples from the Markov Chain “approximate” the posterior distribution;

Posterior distribution

- Assume that 20 patients were treated with 8 responding.
- Prior can be updated into the posterior (again, using MCMC)





Summaries of the posterior distribution

The common summaries of the posterior distribution

- Mean $= 0.405$
- Median $= 0.401$
- Standard deviation $= 0.104$

Posterior means and variances are useful but do not tell us everything.

The posterior distribution can give as various information about the distribution of the parameter.

For example, "*how likely that the response rate is above 25%*":

$$\mathbb{P}(p > 0.25 | \text{Data}, \text{Prior}) = 93.5\%$$

Credible Interval

A specific region which contains a given area of the posterior density function known as a **Credible Interval**.

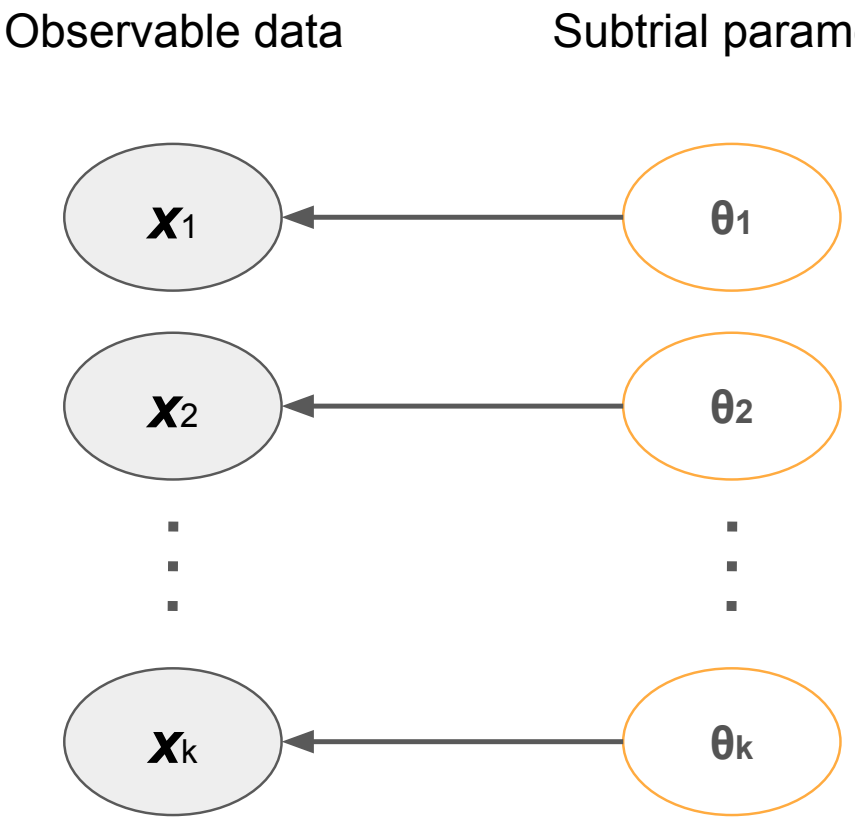
Interpretation: there is a probability of $(1-\alpha) \times 100\%$ that θ falls within the region.

- **Highest density region (HDR)**: smallest interval for θ which contains $(1-\alpha) \times 100\%$ of area.
- **Equal-Tailed Interval:** the interval where the probability of being below the interval is as likely as being above it.

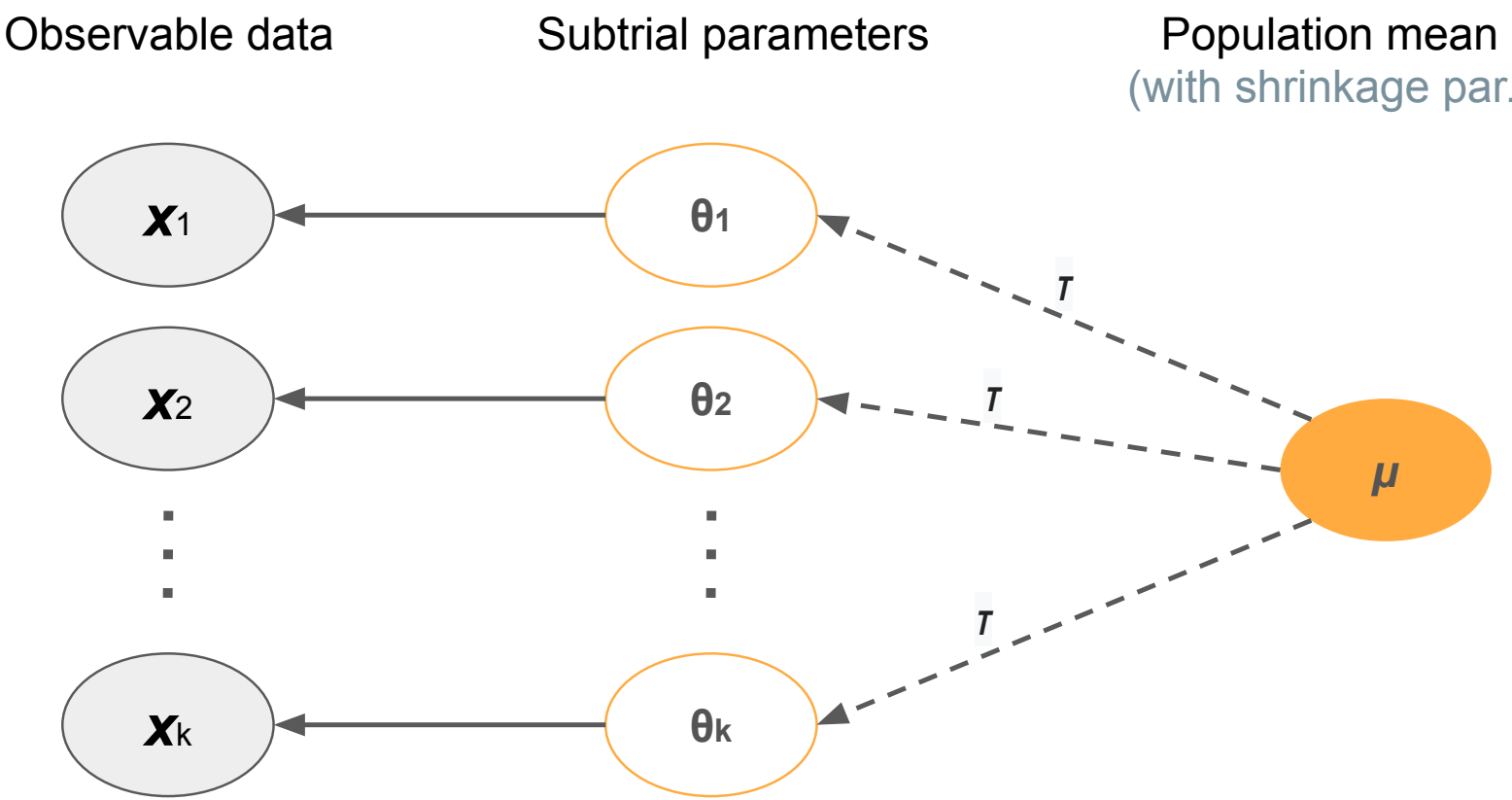
The equal-tailed 95% credible interval in our example is (0.211,0.616)

Bayesian Borrowing Models

Borrowing of information



Borrowing of information



Exchangeability

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 Berry et al. 2013

A Bayesian hierarchical model (BHM) for binomial data:

$$
\begin{aligned}
Y_i &\sim \text{Binomial}(n_i, p_i), && i = 1, \ldots, k \\
\theta_i | \mu, \sigma = \log\left(\frac{p_i}{1-p_i}\right) &\sim N(\mu, \sigma^2), \\
\mu &\sim N(\cdot, \cdot), \quad \sigma \sim g(\cdot).
\end{aligned} \tag{2}
$$

- ⋆ Hierarchical modelling assumes (similarity) of the θ_is
- Borrowing occurs between all baskets $\rightarrow$ the estimates of the response rates are shrunk towards the common mean;
- Degree of shrinkage (i.e. borrowing) controlled by the shrinkage (borrowing) parameter, σ^2.
  - $\sigma^2 = 0 \rightarrow$ complete pooling of data from other baskets;
  - $\sigma^2 = \infty \rightarrow$ no borrowing.

Choice of hyper-prior on borrowing parameter

Based on the uncertainty of θ_j between different baskets, the following classification of values of σ was proposed

- **Small to moderate** heterogeneity: $\sigma = 0.125$ to $\sigma = 0.250$
- **Substantial to large** heterogeneity: $\sigma = 0.5$ to $\sigma = 1$

Various choices of hyper-prior $g(\cdot)$ of σ were proposed with the most common options being

- **Inverse-Gamma**

  Used in the original BHM proposal but was found to lead to a poor behaviour when σ^2 is close to 0 (Cunanan et al. 2019)

- **Half-Cauchy** (with a moderately large scale was suggested instead), e.g. Half-Cauchy$(0, 25)$ (Gelman 2006)
- **Half-Normal** with a prior standard deviation of $s = 0.5$ [95% interval on σ is (0.02, 1.12)] or $s = 1$ [95% interval on σ is (0.03, 2.24)] (Neuenschwander et al. 2016)

Back to the example - I

Now instead of analysing the first basket (8/20) alone, we conduct the analysis together with the second basket (6/18).

We use the BHM with Half-normal prior on σ^2.

The summary characteristics for the first basket:

	Independent	BHM (s = 0.5)	BHM (s = 1.0)
Mean	0.405	0.380	0.385
Median	0.402	0.377	0.381
SD	0.104	0.087	0.091
95% CI	(0.211,0.616)	(0.221,0.559)	(0.218,0.574)
Length of CI	0.404	0.338	0.357

Back to the example - II

Assume that we were less lucky with the second basket and actually the number of responses was 0/18.

We again use the BHM with Half-normal prior on σ^2.

The summary characteristics for the first basket:

	Independent	BHM (s = 0.5)	BHM (s = 1.0)
Mean	0.405	0.321	0.362
Median	0.402	0.314	0.358
SD	0.104	0.102	0.105
95% CI	(0.211,0.616)	(0.145,0.536)	(0.172,0.579)
Length of CI	0.404	0.391	0.407

Notes

- If the baskets are **truly homogeneous**, the gains are higher if smaller variance on the borrowing parameter is assumed
- However, if there is **a heterogeneous basket**, then borrowing can lead to too much shrinkage of the estimates and even increased variance of the estimate
- The choice of the distribution of the borrowing parameters σ^2 should based on the trade-off between these two cases.

Calibrated Bayesian Hierarchical Model Chu et al. 2018

The calibrated BHM, CBHM, has the same form as the BHM, but rather than placing a prior on σ, it defines a measure of homogeneity:

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

where T is the chi-squared test statistic for homogeneity:

$$T = \sum_{i=1}^{k} \frac{(O_{0i} - E_{0i})^2}{E_{0i}} + \sum_{i=1}^{k} \frac{(O_{1i} - E_{1i})^2}{E_{1i}}, \tag{3}$$

where O_{0i} and O_{1i} are the observed failures/responses, while E_{0i} and E_{1i} are the expected failures/responses.

The parameters a and b are tuned via simulations to ensure

- borrowing when all baskets have a homogeneous response
- treat baskets as independent otherwise

Notes on Calibrated BHM

- A benefit of the tuning procedure is the increased certainty in estimates produced by the CBHM in comparison to the BHM in the case where all baskets are homogeneous.

- However, the method takes on a '**strong**' **definition of heterogeneity**: if the response rate in one basket is heterogeneous, then all baskets are deemed heterogeneous, and as a result no borrowing occurs.

- Calibration procedure depends on the chosen scenarios and the sample sizes (!) assumed for these scenarios

Relaxing exchangeability assumption – EXNEX

Neuenschwander et al. 2016

A Bayesian hierarchical model (BHM) assumes that the basket are exchangeable ... with probability of 1(!).

The full exchangeability assumption is often violated. To tackle this, an EXNEX model was proposed

1. EX (exchangeable component): with prior probability w_i, basket i is exchangeable and BHM is applied.
2. NEX (nonexchangeable component): with prior probability $1 - w_i$, θ_i is nonexchangeable with any other basket and analysed independently.

$$
\begin{aligned}
Y_i &\sim \text{Binomial}(n_i, p_i), \\
\theta_i &= \log\left(\frac{p_i}{1-p_i}\right), \\
\theta_i &= \delta_i M_{1i} + (1-\delta_i) M_{2i}, \\
\delta_i &\sim \text{Bernoulli}(w_i),
\end{aligned}
\qquad
\begin{aligned}
M_{1i} &\sim \text{N}(\mu, \sigma^2), \quad \text{(EX)} \\
\mu &\sim \text{N}(\cdot, \cdot), \\
\sigma &\sim g(\cdot), \\
M_{2i} &\sim N(m_i, \nu_i). \quad \text{(NEX)} \quad (4)
\end{aligned}
$$

Back to the example - I

We again analyse the first basket (8/20) together with the second basket (6/18).

We use the EXNEX with Half-normal prior on σ^2 with $s = 0.5$.

The summary characteristics for the first basket:

	Independent	BHM ($s = 0.5$)	EXNEX ($w = 0.5$)
Mean	0.405	0.380	0.395
Median	0.402	0.377	0.390
SD	0.104	0.087	0.097
95% CI	(0.211,0.616)	(0.221,0.559)	(0.216,0.597)
Length of CI	0.404	0.338	0.381

Back to the example - II

Again, assume that we were less lucky with the second basket and actually the number of responses was 0/18.

We use EXNEX with Half-Normal prior on σ^2 with $s = 0.5$.

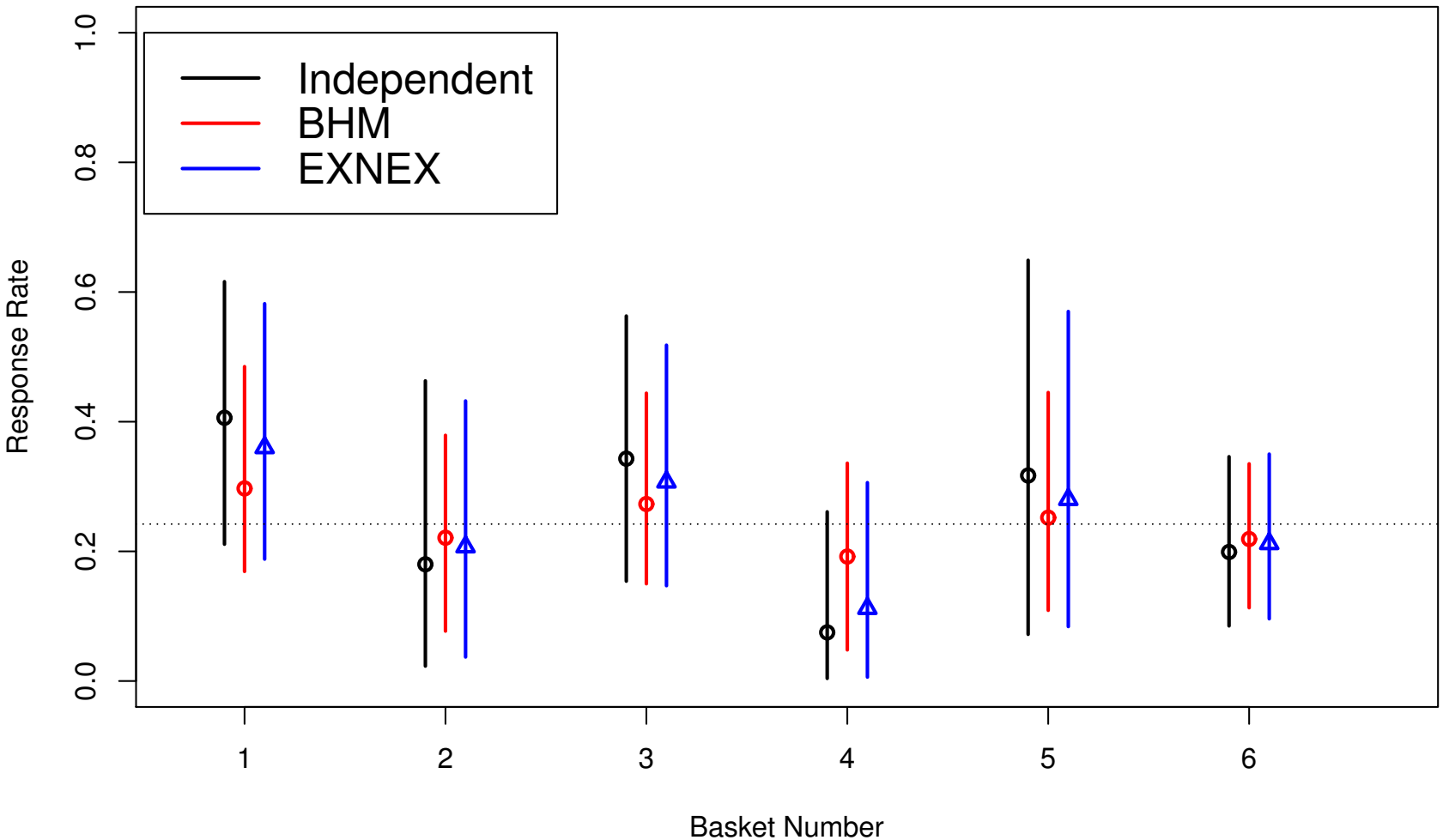
The summary characteristics for the first basket:

	Independent	BHM (s = 0.5)	EXNEX ($w = 0.5$)
Mean	0.405	0.321	0.402
Median	0.402	0.314	0.399
SD	0.104	0.102	0.105
95% CI	(0.211,0.616)	(0.145,0.536)	(0.204,0.614)
Length of CI	0.404	0.391	0.409

Notes on EXNEX

- EXNEX provides more flexibility: information is borrowed only between baskets assigned to the EX component but not from those in the NEX component;
- Weight w_i is the **prior** probability reflecting how likely we believe (a-priori) that the baskets are exchangeable
- It is uncommon to have strong prior information on the probability of exchangeability, so it is suggested to fix $\pi_i = 0.5$
- This prior probability is updated to a some degree based on the homogeneity of the data but...
- The update might not be sensitive enough to the heterogeneity/ homogeneity of responses in the typically small sample sizes

Comparing different analysis models: VE-Basket

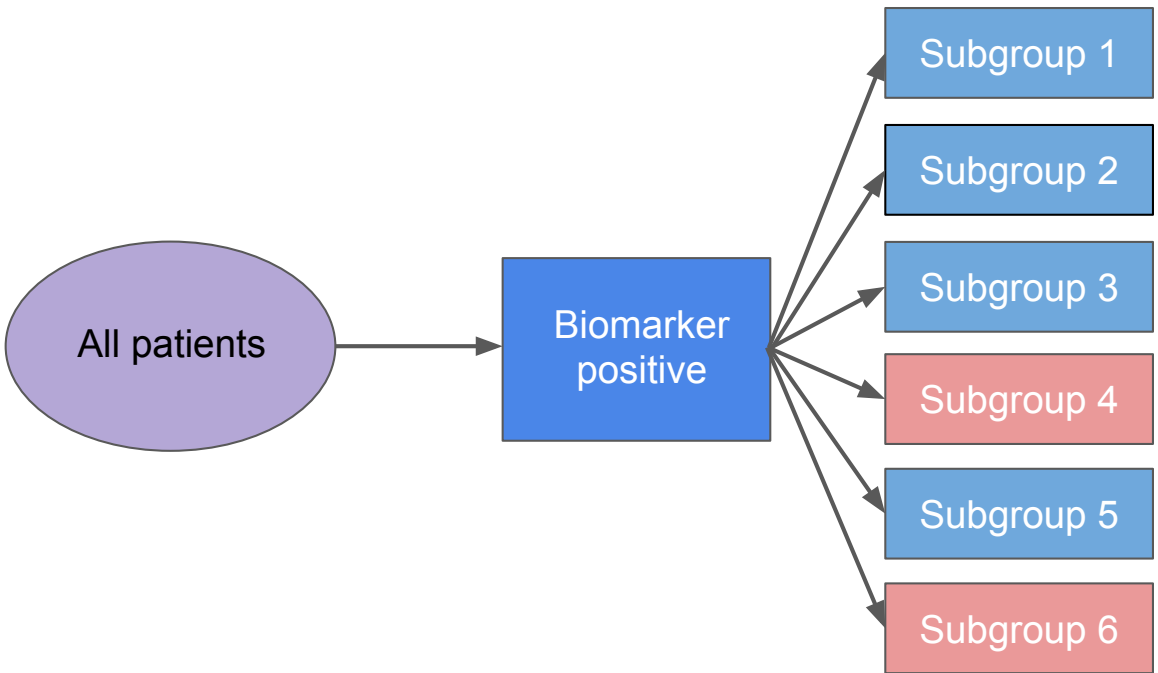


Can we do anything better?

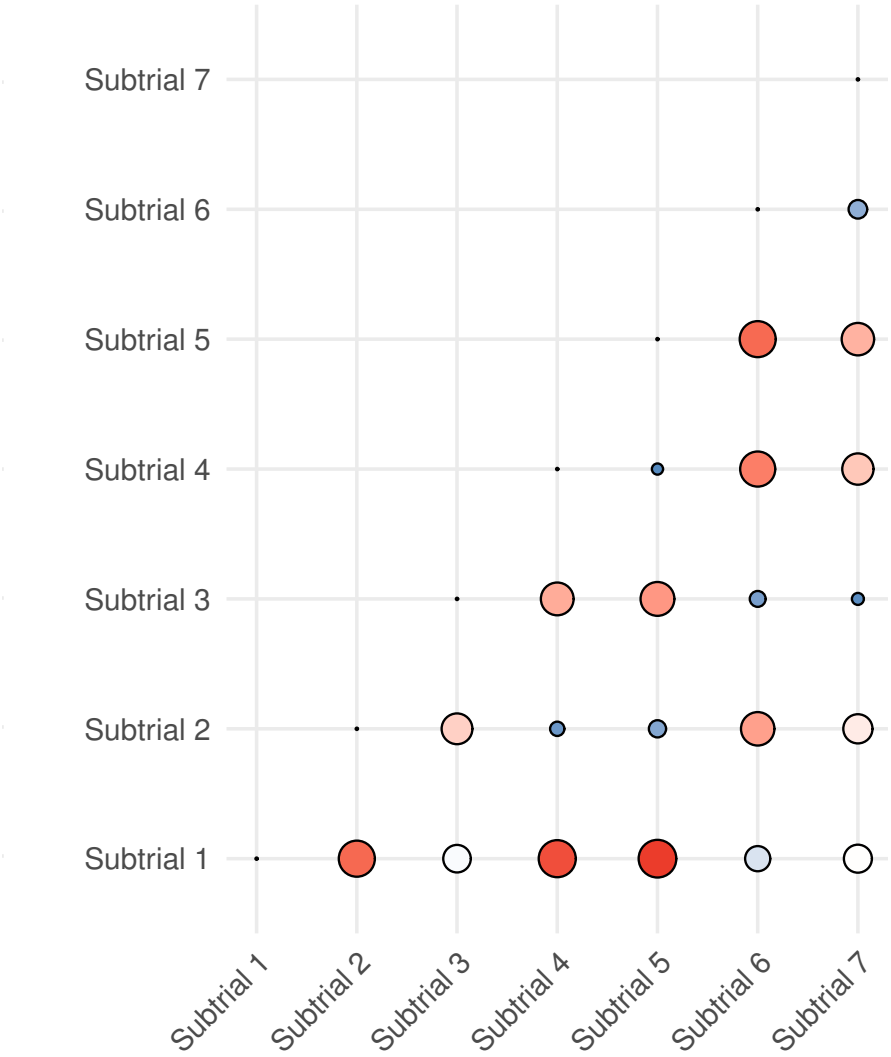
Some modules may display more similar treatment effects between themselves than with others.

Pertinent questions to the EXNEX approach:

- How many EX distributions are needed?
- What is an ‘extreme subgroup’?



Pairwise similarity



Modified EXNEX Daniells et al. (2023)

The mEXNEX$_c$ model is a two-step procedure

- Remove baskets with a clearly heterogeneous response rate

$$\min_{i'}\{\ |\hat{p}_i - \hat{p}_{i'}|\ \} > c, \quad i \neq i',$$

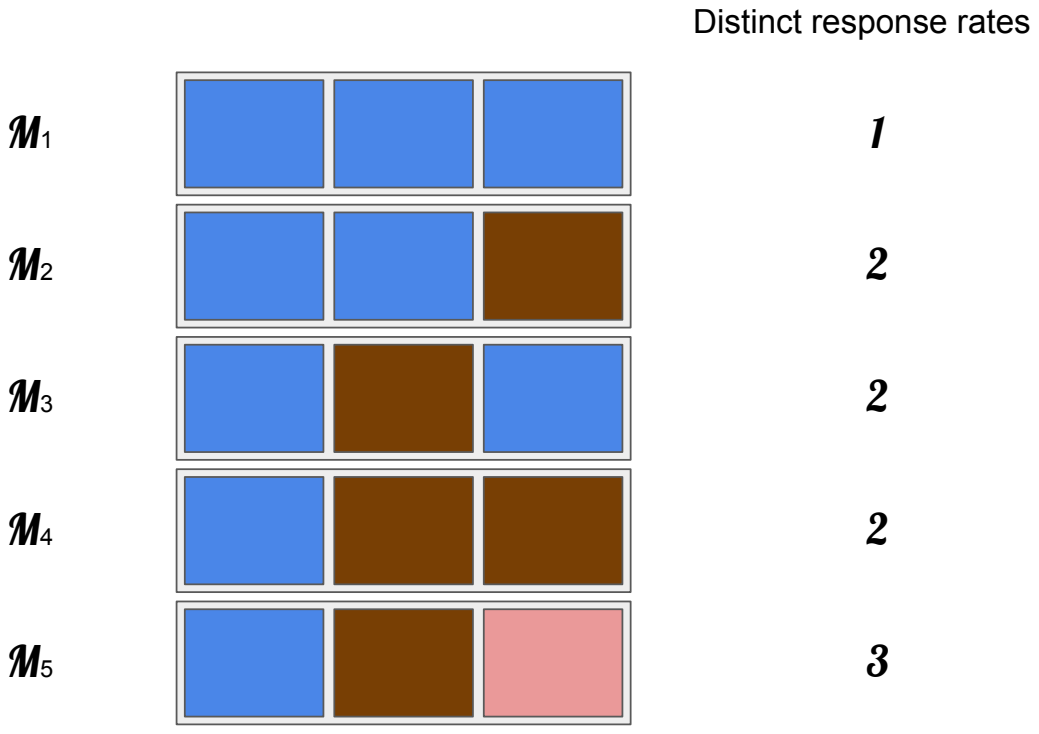
  Treat such a basket as independent and its mixture weight in the EXNEX model is set to 0.

- Find Hellinger pair-wise distance between posteriors (under independence) of basket i and i': $h_{i,i'}$. Compute the weight as average of these distances

$$w_i = \sum_{i'} \frac{1 - h_{i,i'}}{\#\text{ of baskets} - 1} \quad \text{for} \quad i, i' \in S, \quad i \neq i'.$$

- Use these weights in the EXNEX

A model averaging approach Psioda et al. (2019)



- Construct a *complete* model space with all possible reclassification of subgroups with identical p_i
- Each model $\mathcal{M}_\ell$ corresponds to a distinct model stipulation

A model averaging approach Psioda et al. (2019)

Define $\mathcal{M}_j$ as model j representing a permutation of basket allocation to the EX group or NEX group.

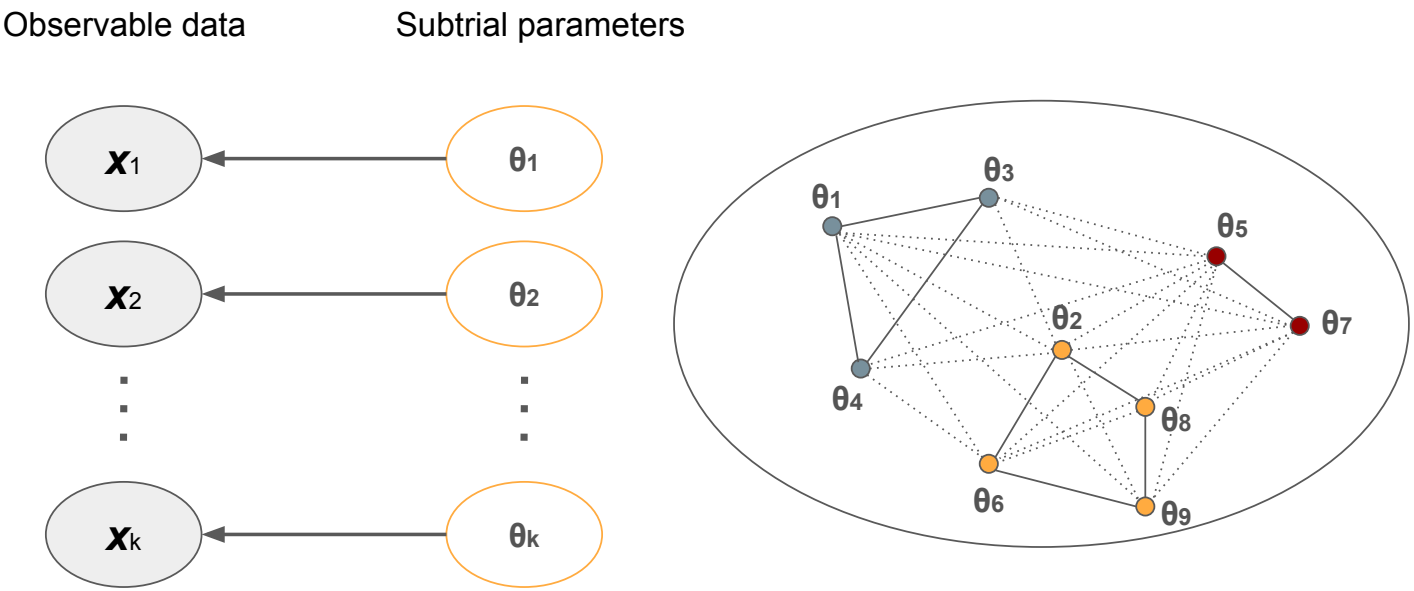
In the EX group, results are pooled and baskets have one shared response rate.

A weakly informative Beta prior is placed on the response rates, while a prior on each model, $f(\mathcal{M}_j)$, is also required. The posteriors $f(p_i|\mathcal{M}_j)$ and $f(\mathcal{M}_j|D)$ are computed.

BMA procedure to obtain a summary statistic for basket k. For example, for the probability of being above 0.25

$$\mathbb{P}(p_i > 0.25|\text{Data},\ \text{Prior}) = \sum_j \mathbb{P}(p_i > 0.25|\mathcal{M}_j, D) f(\mathcal{M}_j|D)$$

Pairwise commensurability Zheng & Wason (2022)



Aim: to estimate the treatment effects, θ_i, using entire trial data.

Robust borrowing of information

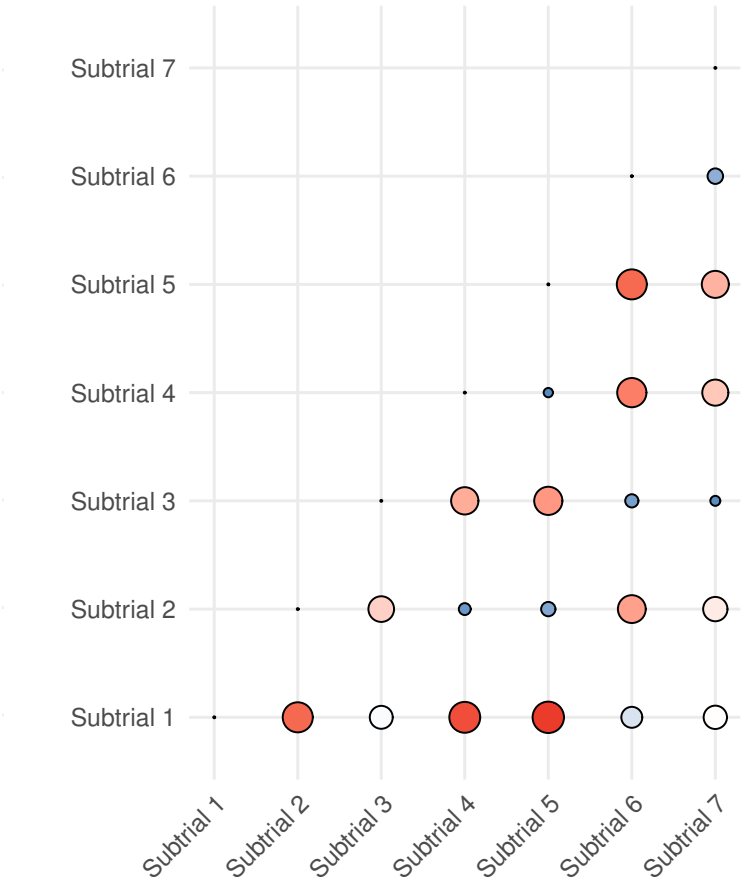
$$\theta_i \mid \theta_q, \nu_{qi} \sim N(\theta_q, \nu_{qi}^{-1}), \forall i = 1, \ldots, k$$

$$\nu_{qi} \sim w_{qi}\text{Gamma}(a_1, b_1) +$$

$$(1 - w_{qi})\text{Gamma}(a_2, b_2), \text{ with } q \neq i$$

- Gamma(a_1, b_1) correspond to substantial borrowing
- Gamma(a_2, b_2) correspond to limited borrowing
- Setting $w_{qi} \to 0$ means strong borrowing and 1 means no borrowing.

Robust borrowing of information



- w_{qi} measures dissimilarity
- e.g. can be based on Hellinger Distance

$$w_{qi} = \left[1 - \sqrt{\frac{2\sigma_q\sigma_i}{\sigma_q^2+\sigma_i^2}}\exp\left(-\frac{(\mu_q-\mu_i)^2}{4(\sigma_q^2+\sigma_i^2)}\right)\right]^{1/2} \quad \text{for } Y_i \sim N(\mu_i, \sigma_i^2)$$

Obtaining a collective prior

Given all pairwise commensurate priors $\pi(\theta_i|\boldsymbol{y}_q)$, $\forall i \neq q$, we then synthesise $K-1$ commensurate predictive priors.

The synthesis weights are

$$d_{qi} = \frac{\exp(-w_{qi}^2/r_0)}{\sum_q \exp(-w_{qi}^2/r_0)}$$

where $\sum_q d_{qi} = 1$ and r_0 is the parameters that governs how much influence the Hellinger distance has on the weight.

The resulting collective commensurate predictive prior

$$\theta_i|\boldsymbol{y}_{-i} \sim \mathcal{N}\left(\sum_{q\neq i} d_{qi}\lambda_q, \sum_{q\neq i} d_{qi}^2\xi_{qi}^2\right)$$

where λ_q is the mean of $\theta_q|\boldsymbol{y}_q$ and ξ_{qi}^2 is the variance of $\theta_i|\boldsymbol{y}_q$

How to design a basket trial

Decision-making under Bayesian inference

- We have covered various analysis methods, but how one can plan for such a trial?
  - What would be the type I error? Power?
    If (some of) the baskets are not homogeneous, one can expect an inflation of type I or/and decrease in power
  - What are the properties of the estimation?
- A conventional approach is to **fix the sample size** (assuming no borrowing or just based on feasibility) and then conduct simulation studies to evaluate the operating characteristics
- An alternative would be to actually **plan for the study** (i.e. for the sample size) **under the assumption of borrowing**
- We will cover both approaches

Bayesian decision-making

- Decision are made based on the **posterior probabilities for the treatment effect** given the clinical trial data;

For example, in a single-arm trial, one is claiming **benefit** if

$$\text{Prob}\left[\theta \geq p_0 \mid \text{Data},\ \text{Prior}\right] > c$$

where

- p_0 is an effect threshold
- c is probability threshold (often chosen to control type I error)
- ‘Data” is the observed trial data
- “Prior” is specified prior distribution on the treatment effect.

Bayesian decision-making: example

- Assume that the planned sample size is $N = 13$ patients;
- The null response rate is $p_0 = 0.15$
- Clinically interesting response rate is $p_1 = 0.45$

The type I error is defined as

$$\text{Prob}\{\text{Prob}\left[\theta \geq p_0 \mid \text{Data},\ \text{Prior}\right] > c | \theta = p_0\}$$

The power is defined as

$$\text{Prob}\{\text{Prob}\left[\theta \geq p_0 \mid \text{Data},\ \text{Prior}\right] > c | \theta = p_1\}$$

Scenarios to consider

For a study with 5 baskets:

	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	0.15	0.15	0.15
Scenario 4	0.45	0.45	0.45	0.15	0.15
Scenario 5	0.45	0.45	0.45	0.45	0.15
Scenario 6	0.45	0.45	0.45	0.45	0.45

Scenarios to consider

For a study with 5 baskets:

	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	0.15	0.15	0.15
Scenario 4	0.45	0.45	0.45	0.15	0.15
Scenario 5	0.45	0.45	0.45	0.45	0.15
Scenario 6	0.45	0.45	0.45	0.45	0.45

Equal sample sizes?

Scenarios to consider

For a study with 5 baskets:

	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	0.15	0.15	0.15
Scenario 4	0.45	0.45	0.45	0.15	0.15
Scenario 5	0.45	0.45	0.45	0.45	0.15
Scenario 6	0.45	0.45	0.45	0.45	0.45

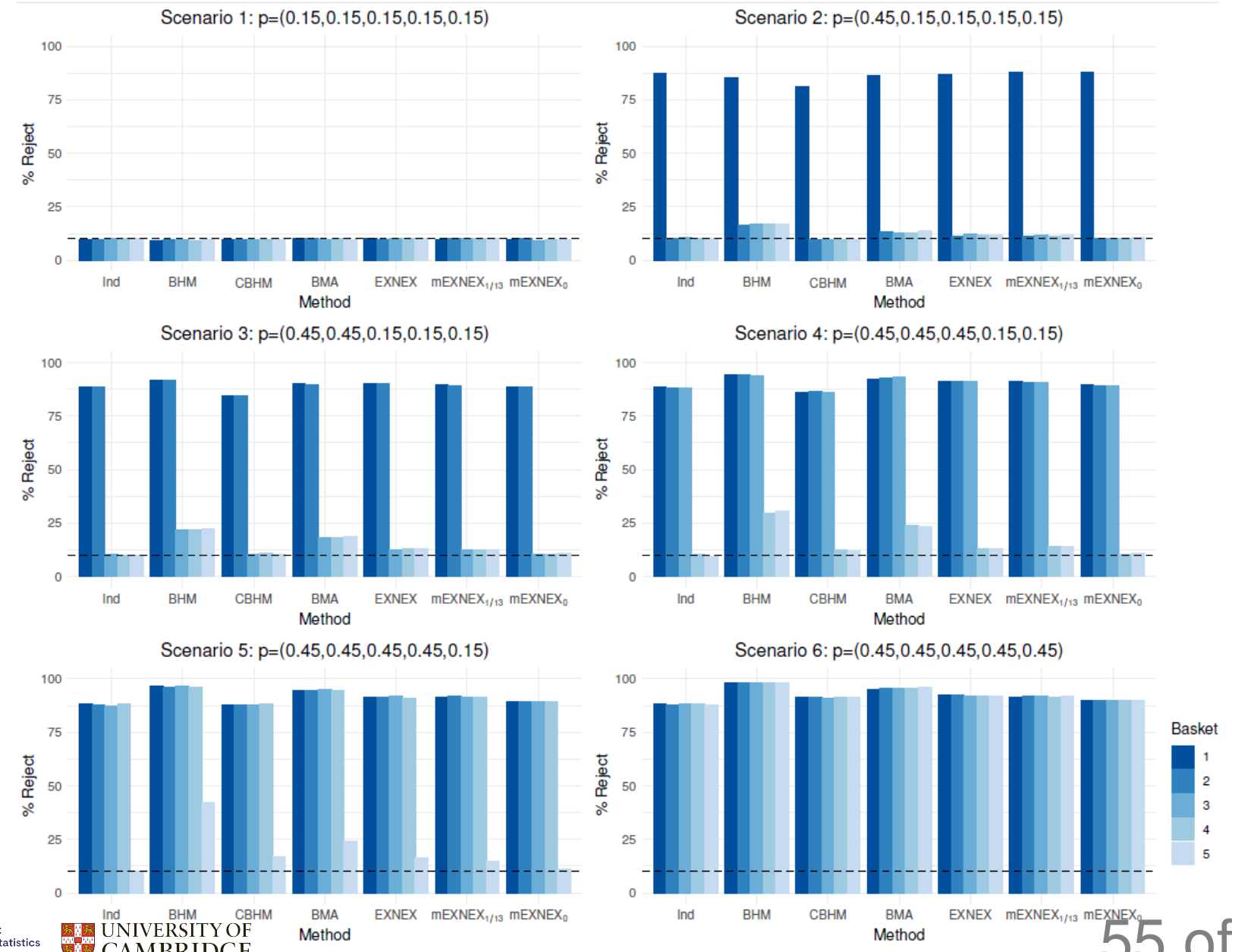
Equal sample sizes?

Moderate treatment effect?

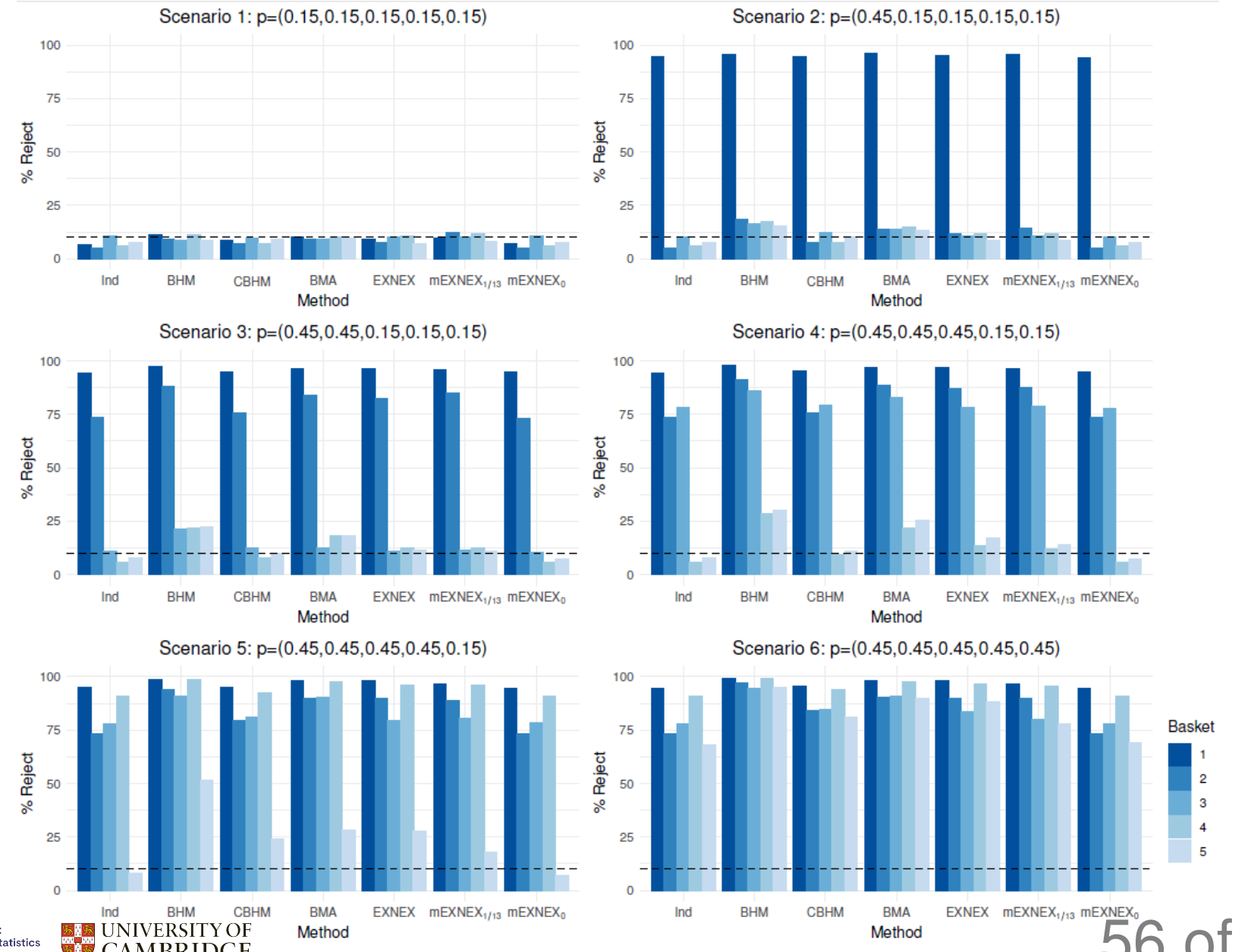
Null configuration

- What would be the null hypothesis in a basket trial setting?
- With K baskets, one can consider the configuration $p_1 = p_2 = \ldots = p_k = p_0$ where p_0 is a null response rate for the calibration of the probability threshold c.
- This assumes that all baskets are homogeneous. What if some of the baskets has interesting treatment effect and some null?
- Were one to calibrate under the case of the maximum type I error, there will be no more advantages of using borrowing (Kopp-Schneider et al. 2020)

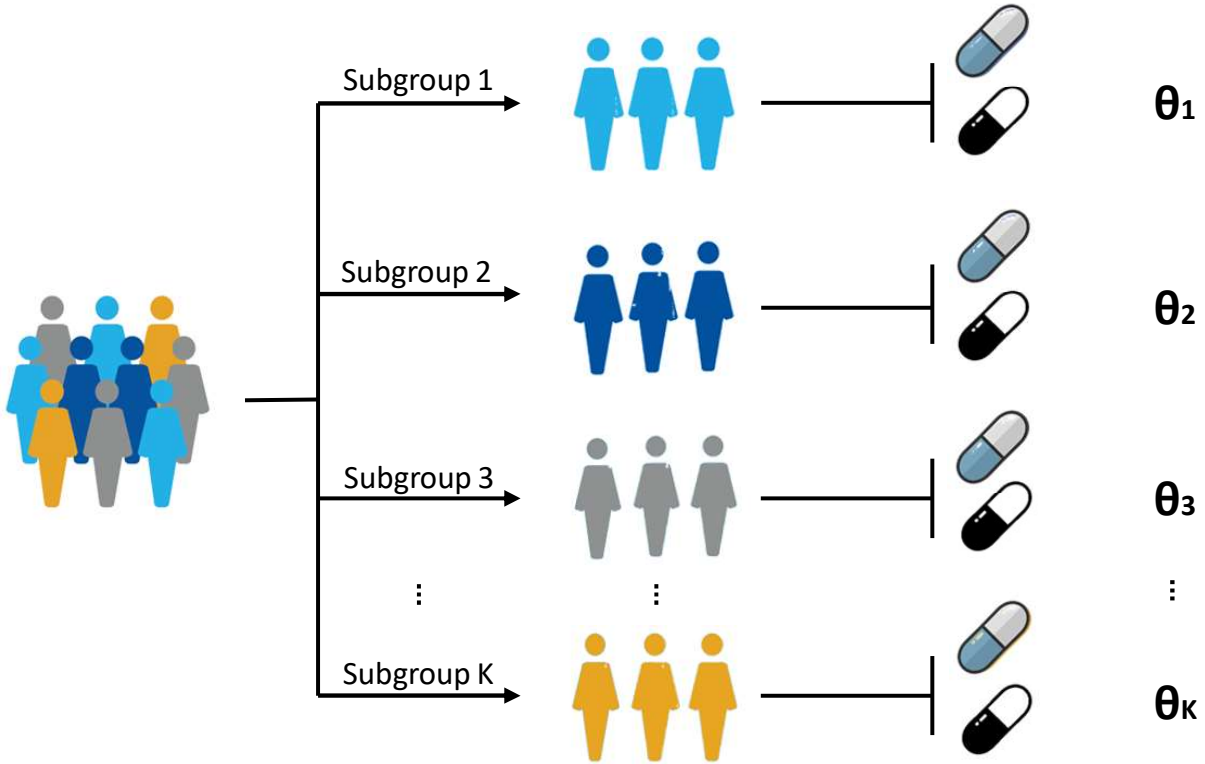
Results: Planned sample size Daniells et al. (2023)



Results: Realised sample size (as in practical)



Randomised controlled basket trials



Randomised controlled basket trials

In a single-arm setting:

$$Y_i \mid p_i, n_i \sim \text{Binomial}(p_i, n_i)$$
$$\text{logit}(p_i) = \theta_i$$

In a randomised controlled setting:

$$Y_{jk} \mid p_{ji}, n_{ji} \sim \text{Binomial}(p_{ji}, n_{ji}),\ j = E, C,$$
$$\text{logit}(p_{Ci}) = \alpha_i$$
$$\text{logit}(p_{Ei}) = \alpha_i + \theta_i$$

where an uninformative prior is placed on α_i, while $\theta_i \mid \mu, \tau \sim N(\mu, \tau^2)$ to implement borrowing.

To observe n patients.

Patients are randomised to receive the new treatment (w.p. R) or the control (w.p. $1 - R$).

Let $\bar{X}_i$ be the mean response per treatment group $i = E, C$, so that

$$\bar{X}_E - \bar{X}_C \sim N\left(\theta, \frac{\sigma^2}{nR(1-R)}\right), \quad \text{with } \theta = \mu_E - \mu_C.$$

Assume known σ^2.

Specify a prior on the difference in means, $\theta \sim N(m, s^2)$.

Establish **efficacy** if $\Pr(\theta \geq 0 \mid \text{data, prior}) \geq \eta$, or
futility if $\Pr(\theta < \delta \mid \text{data, prior}) \geq \zeta$.

Sample size formula under no borrowing

Under the assumption of no borrowing, the sample size required is

$$n \geq \frac{\sigma^2}{R(1-R)}\left[\left(\frac{z_\eta + z_\zeta}{\delta}\right)^2 - \frac{1}{s^2}\right].$$

Sample size formula under commensurate prior

Zheng et al. (2023)

Recall

$$\theta_i \mid \theta_q, \nu_{qi} \sim N(\theta_q, \nu_{qi}^{-1}),\ \forall i = 1, \ldots, k$$

$$\nu_{qi} \sim w_{qi}\text{Gamma}(a_1, b_1) + (1 - w_{qi})\text{Gamma}(a_2, b_2),\ \text{with } q \neq i$$

- Setting $w_{qi} \to 0$ means strong borrowing

For basket-specific sample sizes $n_1, \ldots, n_k$ satisfy

$$\frac{R_k(1-R_k)n_k}{\sigma_k^2}\left[\sum_q d_{qk}^2\left(\left(\frac{1}{s_{0q}^2} + \frac{R_q(1-R_q)n_q}{\sigma_q^2}\right)^{-1} + \frac{w_{qk}b_1}{a_1 - 1} + \frac{(1-w_{qk})b_2}{a_2 - 1}\right)\right]^{-1} \geq \left(\frac{z_\eta + z_\zeta}{\delta}\right)^2 \ \forall k \neq q$$

- With $0 \leq w_{qk} \leq 1$, sample size saving can be expected

Apply Newton's method for systems of nonlinear equations to find $n_1, \ldots, n_k$, simultaneously

If you are a little rusty on Newton's method...

Sample size formulae for cases of no borrowing and commensurat borrowing are implemented in an online `R` Shiny App

`http://shiny.mrc-bsu.cam.ac.uk/apps/BasketTrials/.`

R Shiny App: Sample Size Formula

# of Subtrials	*outcome variances* σ^2_1	a_1	*outcome means* μ_1
5	10	2	-0.489
δ: effect size	σ^2_2	b_1	μ_2
0.8	10	2	0.226
η: efficacy threshold	σ^2_3	a_2	μ_3
0.95	10	54	-0.181
ζ: futility threshold	σ^2_4	b_2	μ_4
0.8	10	3	0.293
	σ^2_5		μ_5
	10		0.329

Enable borrowing:

Level of Borrowing:

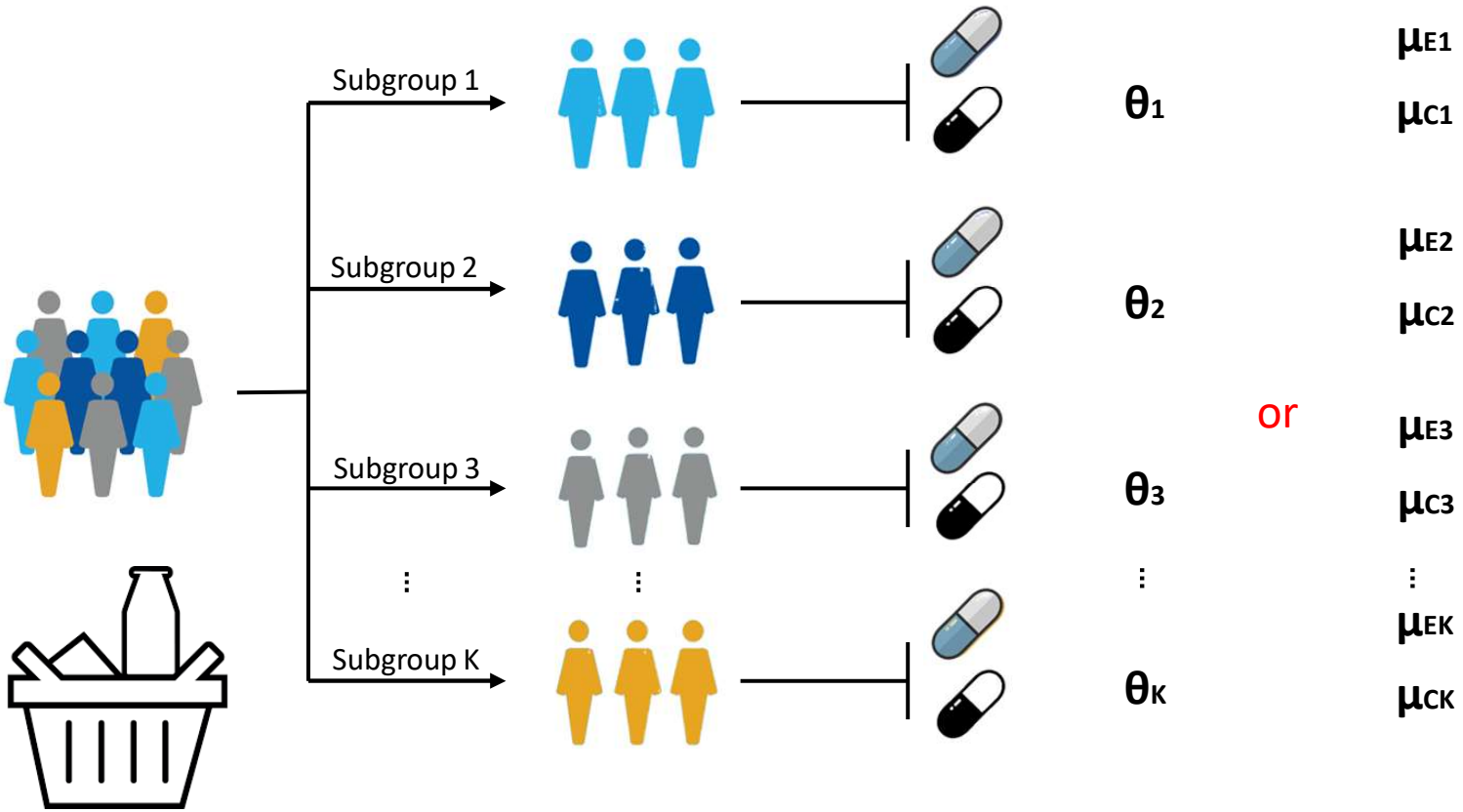
○ Moderate ○ Strong ◉ Pairwise

Use custom allocation ratios:

Results:

k	σ_k	n	n_E	n_C
1	10	136	68	68
2	10	78	39	39

An alternative strategy for borrowing Ouma et al. (2022)



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
- If accounting for this formally in the design stage, borrowing means a smaller sample size to reach the same desired level of power/decision accuracy
- Inflation of type I error rates under BHM (Freidlin and Korn, 2013) and other borrowing methods

References (1)

Berry S, Broglio K, Groshen S, Berry D. (2013). Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. Clinical Trials.

Chu Y, Yuan Y. A Bayesian basket trial design using a calibrated Bayesian hierarchical model. Clinical Trials. 2018 Apr;15(2):149-58.

Cummings J, Montes A, Kamboj S, Cacho JF. (2022) The role of basket trials in drug development for neurodegenerative disorders. Alzheimer's Research and Therapy.

Cunanan, K, Iasonos A, Shen R, Gönen M. (2019). Variance prior specification for a basket trial design using Bayesian hierarchical modeling. Clinical Trials.

Daniells L, Mozgunov P, 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.

Fouarge E, Monseur A, Boulanger B, et al. (2021) Hierarchical Bayesian modelling of disease progression to inform clinical trial design in centronuclear myopathy. Orphanet Journal of Rare Diseases.

Freidlin B, Korn EL. (2013) Borrowing Information across Subgroups in Phase II Trials: Is It Useful? Borrowing Information across Subgroups. Clinical Cancer Research.

Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). (2006) Bayesian Analysis

Hyman D, Puzanov I, Subbiah V, et al. (2017) Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. NEJM

References (2)

Kopp-Schneider A, Calderazzo S, Wiesenfarth M.(2020) Power gains by using external information in clinical trials are typically not possible when requiring strict type I error control. Biometrical Journal.

Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. (2016) Robust exchangeability designs for early phase clinical trials with multiple strata. Pharmaceutical Statistics

Ouma L, Grayling M, Wason JMS, Zheng H. (2022) Bayesian modelling strategies for borrowing of information in randomised basket trials. JRSS-C

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

Whitehead J, Valdés-Márquez E, Johnson P, Graham G. (2008) Bayesian sample size for exploratory clinical trials incorporating historical data. Statistics in Medicine.

Zheng H, Grayling MJ, Mozgunov P, Jaki T, Wason JMS. (2023) Bayesian sample size determination in basket trials borrowing information between subsets. Biostatistics.

Zheng H, Wason JMS. (2022). Borrowing of information across patient subgroups in a basket trial based on distributional discrepancy. Biostatistics.