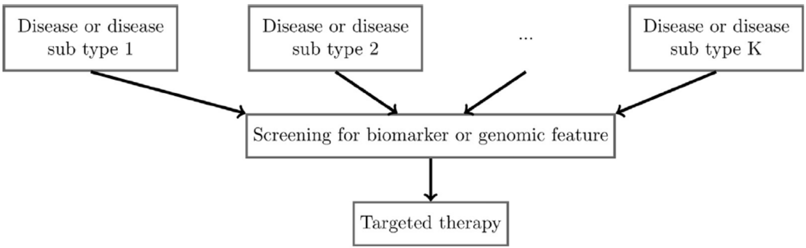
Bayesian Basket Trial Design with False Discovery Rate Control

Emily C. Zabor
Department of Quantitative Health Sciences
Taussig Cancer Institute

IISA conference
May 20, 2021

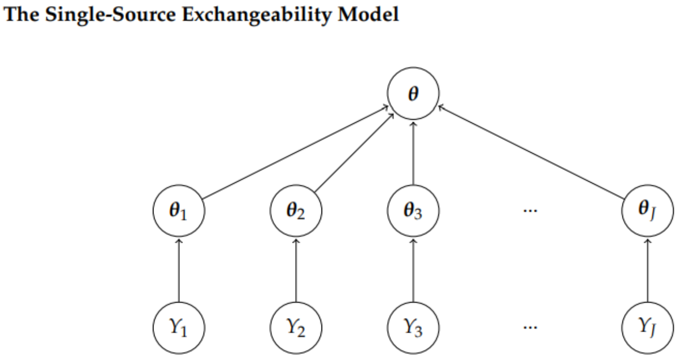
Cleveland Clinic

Basket trials are becoming increasingly common in oncology



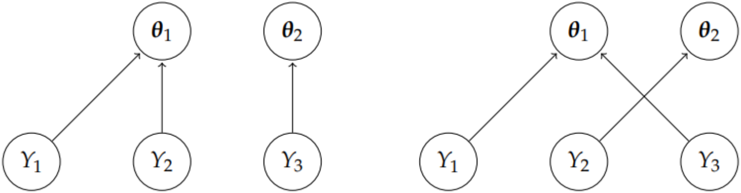
Meyer et al. 2020. The Evolution of Master Protocol Clinical Trial Designs: A Systematic Literature Review. Clinical Therapeutics.

Single-source exchangeability models average everyone



Kane et al. 2020. Analyzing Basket Trials under Multisource Exchangeability Assumptions. The R Journal Vol 12/2.

Multi-source exchangeability models can enumerate all possible exchangeability configurations



(a) Model where Y_1 and Y_2 are exchangeable. **(b)** Model where Y_1 and Y_3 are exchangeable.

Kane et al. 2020. Analyzing Basket Trials under Multisource Exchangeability Assumptions. The R Journal Vol 12/2.

Multi-source exchangeability model details

Received: 11 August 2017 | Revised: 17 March 2018 | Accepted: 8 June 2018

DOI: 10.1002/sim.7893

RESEARCH ARTICLE

WILEY Statistics in Medicine

Bayesian basket trial design with exchangeability monitoring

Brian P. Hobbs[1] | Rick Landin[2]

Hobbs and Landin. 2018. Bayesian basket trial design with exchangeability monitoring. Statistics in Medicine 37(25): 3557-3572.

Analyzing MEMs with the basket package in R

Two fitting options:

1. `mem_mcmc`: Bayesian Metropolis-Hasting MCMC inference
2. `mem_exact`: Full Bayesian inference

Method	Return Description
`basket_pep`	Basketwise PEP matrix
`basket_map`	Basketwise MAP matrix

Plot Method	Return Description
`plot_pep_graph`	Network graph of the PEP matrix
`plot_pep`	Exchangeogram of the PEP matrix
`plot_map`	Exchangeogram of the MAP matrix

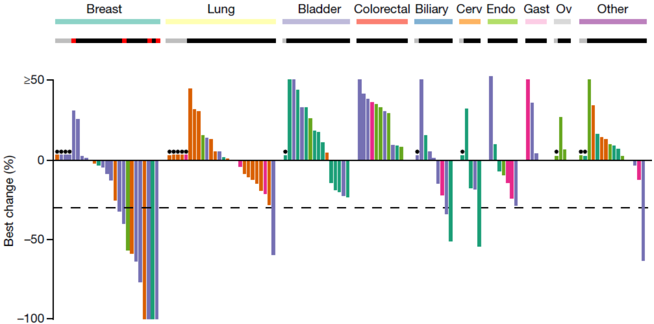
Kane et al. 2020. Analyzing Basket Trials under Multisource Exchangeability Assumptions. The R Journal Vol 12/2.

Case study: the SUMMIT trial design

- Testing neratinib in HER2- and HER3-mutant tumors
- ORR $\leq 10\%$ unacceptable, ORR $\geq 30\%$ acceptable
- Independent Simon's optimal two-stage designs
  - Enroll 7 patients
  - If at least 1 response, enroll additional 11 patients
  - If 4 responses seen in 18 total patients, reject null

Hyman et al. (2018). Her kinase inhibition in patients with her2- and her3-mutant cancers. Nature 554(7691), 189-194.

Case study: the SUMMIT trial results

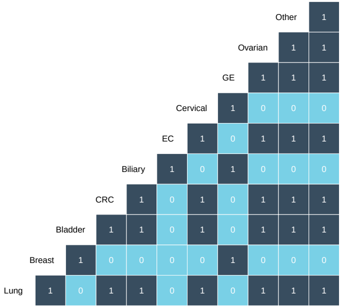


- Breast ORR 32%

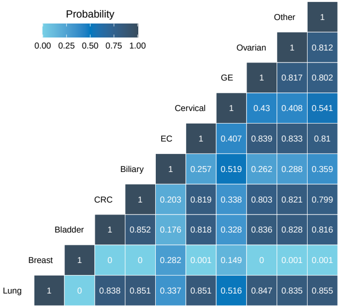
Hyman et al. (2018). Her kinase inhibition in patients with her2- and her3-mutant cancers. Nature 554(7691), 189-194.

Case study: exchangeograms of SUMMIT MAP and PEP

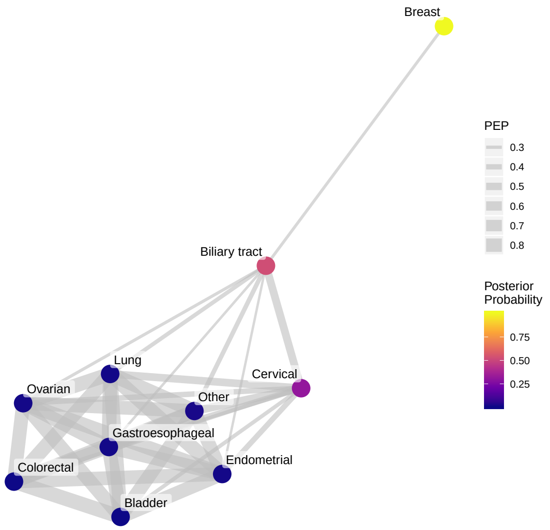
Maximum A Posteriori MEM



Posterior Exchangeability Probability



Network graph of SUMMIT results



Cleveland Clinic

The issue of multiplicity

https://research.bigagainstbreastcancer.org/research/translational-research

False Discovery Rate control

1. Obtain the posterior probability that each basket j exceeds the null response rate, $\Pr(\pi_j > \pi_0 | \boldsymbol{S})$
2. Order the posterior probabilities from largest to smallest, $\Pr(\pi_j > \pi_0 | \boldsymbol{S})_{(1)}, \ldots, \Pr(\pi_j > \pi_0 | \boldsymbol{S})_{(j)}$
3. For a given threshold of posterior probability, ϕ, identify the largest k such that $\Pr(\pi_j > \pi_0 | \boldsymbol{S})_{(k)} > \frac{k}{j} \times \phi$
4. Declare all baskets with posterior probability $\Pr(\pi_j > \pi_0 | \boldsymbol{S})_{(i)}$, $i = 1, \ldots, k$ to be significant at threshold ϕ

Benjamini et al. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57(1), 289-300.

MEM with FDR control simulation design

Table 1. True response probabilities used to compare trial operating characteristics between Bayesian design with MEM and independent frequentist design.

	Basket									
	Lung	Breast	Bladder	CRC	Biliary	EC	Cervical	GE	Ovarian	Other
Scenario	n=26	n=25	n=18	n=17	n=11	n=8	n=5	n=7	n=5	n=19
Global Alt	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Mixed Alt 1	0.10	0.30	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Mixed Alt 2	0.10	0.30	0.10	0.10	0.30	0.10	0.10	0.10	0.10	0.10
Mixed Alt 3	0.10	0.30	0.10	0.10	0.30	0.10	0.30	0.10	0.10	0.10
Global Null	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10

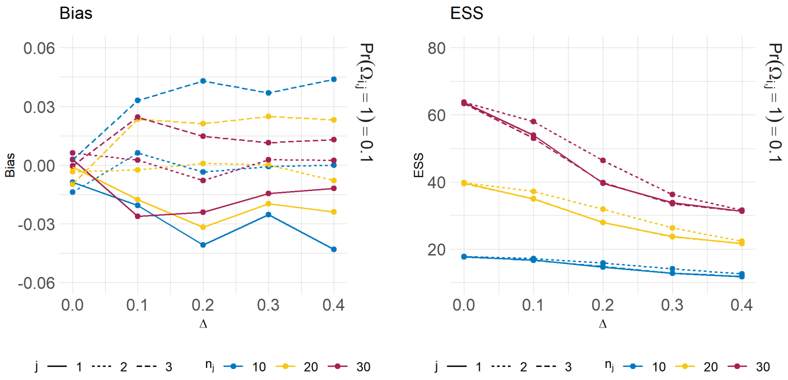
Alt = Alternative; CRC=Colorectal; EC=Endometrial; GE=Gastroesophageal

Simulation to investigate specification of the prior probability of exchangeability

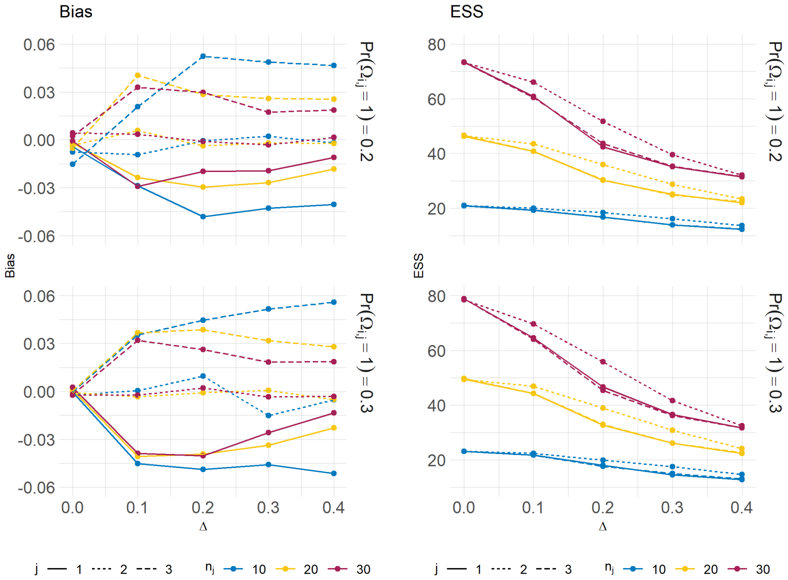
	j=1	j=2	j=3
j=1	1	$Pr(\Omega_{i,j} = 1)$	$Pr(\Omega_{i,j} = 1)$
j=2	$Pr(\Omega_{i,j} = 1)$	1	$Pr(\Omega_{i,j} = 1)$
j=3	$Pr(\Omega_{i,j} = 1)$	$Pr(\Omega_{i,j} = 1)$	1

- Response probabilities for the 3 baskets: $(0.5, 0.5, 0.5)$, $(0.4, 0.5, 0.6)$, $(0.3, 0.5, 0.7)$, $(0.2, 0.5, 0.8)$, $(0.1, 0.5, 0.9)$
- Δ is absolute difference between $\pi_j, j \in (1, 3)$, and π_2
- Sample size equal across baskets $n_j = 10, 20, 30$
- Null response rate fixed at 0.15

Calibrating the prior probability of exchangeability



Calibrating the prior probability of exchangeability



Cleveland Clinic

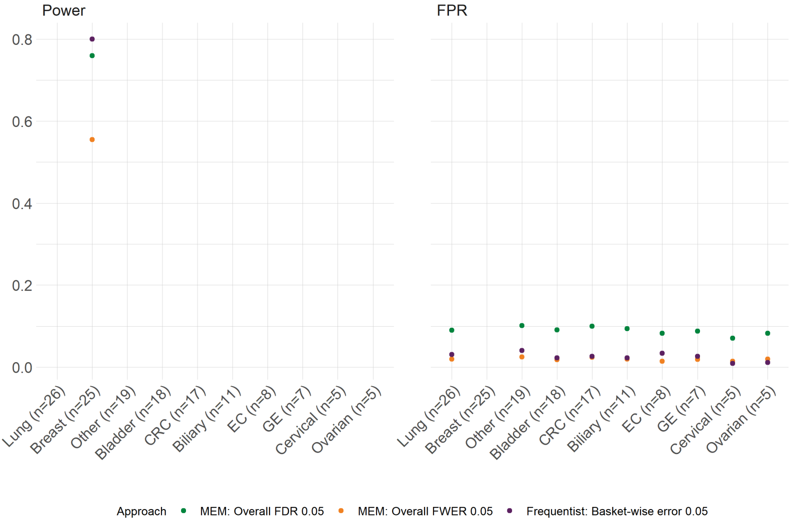
Overall results using the MEM approach

- Posterior threshold 0.867 controls the FDR at 0.1
- Posterior threshold 0.946 controls the FDR at 0.05

Table 2. Power, false positive rate (FPR) and family-wise error rate (FWER) for each scenario using the MEM approach.

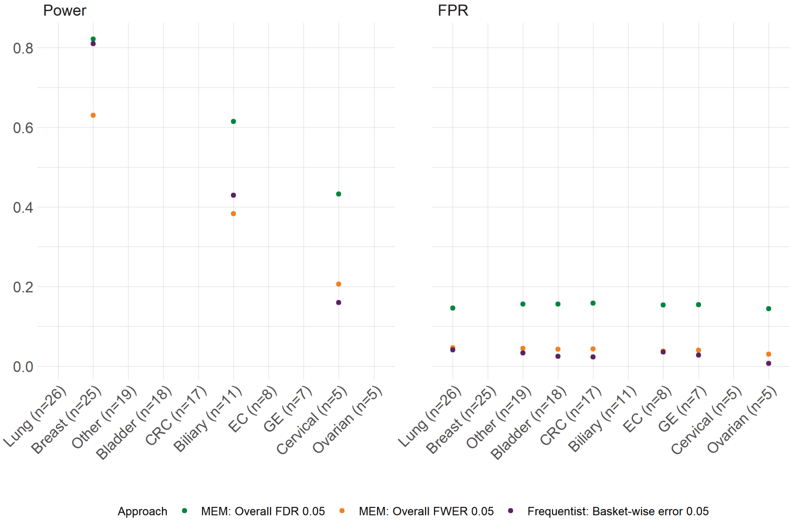
	FDR = 0.1			FDR = 0.05			FWER = 0.05	
Setting	Power	FPR	FWER	Power	FPR	FWER	Power	FPR
Global Null	–	0.10	0.32	–	0.04	0.17	–	0.01
Mixed Alternative 1	0.86	0.18	0.52	0.76	0.09	0.31	0.56	0.02
Mixed Alternative 2	0.81	0.24	0.58	0.69	0.12	0.38	0.48	0.03
Mixed Alternative 3	0.75	0.27	0.58	0.62	0.15	0.41	0.41	0.04
Global Alternative	0.96	–	–	0.94	–	–	0.90	–

Basket-wise results: mixed alternative 1



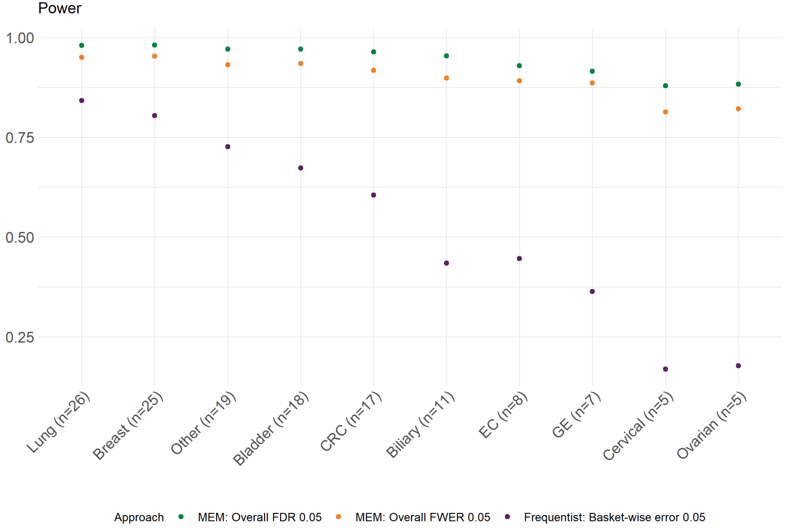
Cleveland Clinic

Basket-wise results: mixed alternative 3



Cleveland Clinic

Basket-wise results: global alternative



Summary

- MEM information-sharing can compensate for small basket sample sizes and improve power
- MEM provides a direct framework for examining heterogeneity across baskets through the posterior exchangeability probabilities
- FDR control maintains higher power at the cost of higher FPR
- Easy implementation through the `basket` package in `R`

Questions?