

# **Bayesian Modeling in Oncology Trials**

*Part 2: Information Sharing in Clinical Trials*

Stat4Onc Short Course

May 10, 2023

# Center for Innovative Design & Analysis

colorado school of public health

---

## Instructors



### **Dr. Brian Hobbs**

Associate Professor of Biostatistics, The University of Texas Dell Medical School

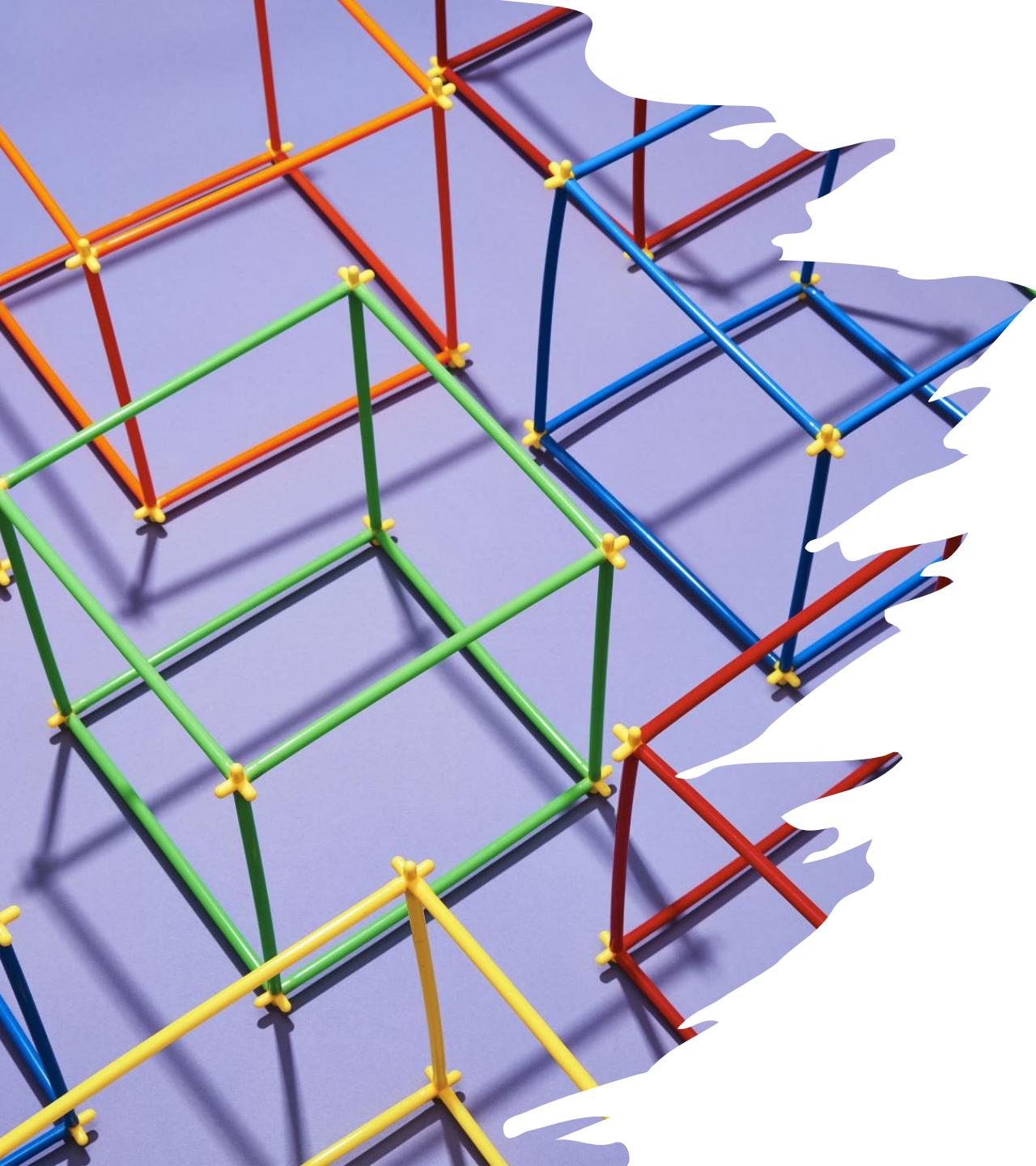
Email: [brian.hobbs@austin.utexas.edu](mailto:brian.hobbs@austin.utexas.edu)



### **Dr. Alex Kaiser**

Assistant Professor of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus

Email: [alex.kaiser@cuanschutz.edu](mailto:alex.kaiser@cuanschutz.edu)



# Short Course Outline

1. Introduction to Bayesian Modeling in Oncology Trials
2. Information Sharing in Clinical Trials
3. Statistical Considerations for Master Protocols
4. Basket Trial Software and Examples

# Overall Learning Objectives

1. Exposure to the basics of Bayesian methods in oncology trials
2. Introduction of methods for sharing information across data sources
3. Understand use of master protocol studies (e.g., platform and basket trials)
4. Bayesian models for hybrid randomized controlled trials

# Precision Oncology Design

ACCEPTED MANUSCRIPT

## Statistical challenges posed by uncontrolled master protocols: sensitivity analysis of the Vemurafenib study

B P Hobbs , M J Kane, D S Hong, R Landin

Annals of Oncology, mdy457, <https://doi.org/10.1093/annonc/mdi457>

Published: 18 October 2018

ANNALS  
OF  
ONCOLOGY

Split View PDF Cite Permissions Share ▾

### Abstract

Within the evidentiary hierarchy of experimental inquiry, randomized trials are the gold standard. Oncology patients enter clinical studies with diverse lifestyles, treatment pathways, host tissue

review articles

Progress in the areas of genomics, disease pathways, and drug discovery has advanced into clinical and translational cancer research. The latest innovations in clinical trials have followed with master protocols, which are defined by inclusive eligibility criteria and devised to interrogate multiple therapies for a given tumor histology and/or multiple histologies for a given therapy under one protocol. The use of master protocols for oncology has become more common with the desire to improve the efficiency of clinical research and accelerate overall drug development. Basket trials have been devised to ascertain the extent to which a treatment strategy offers benefit

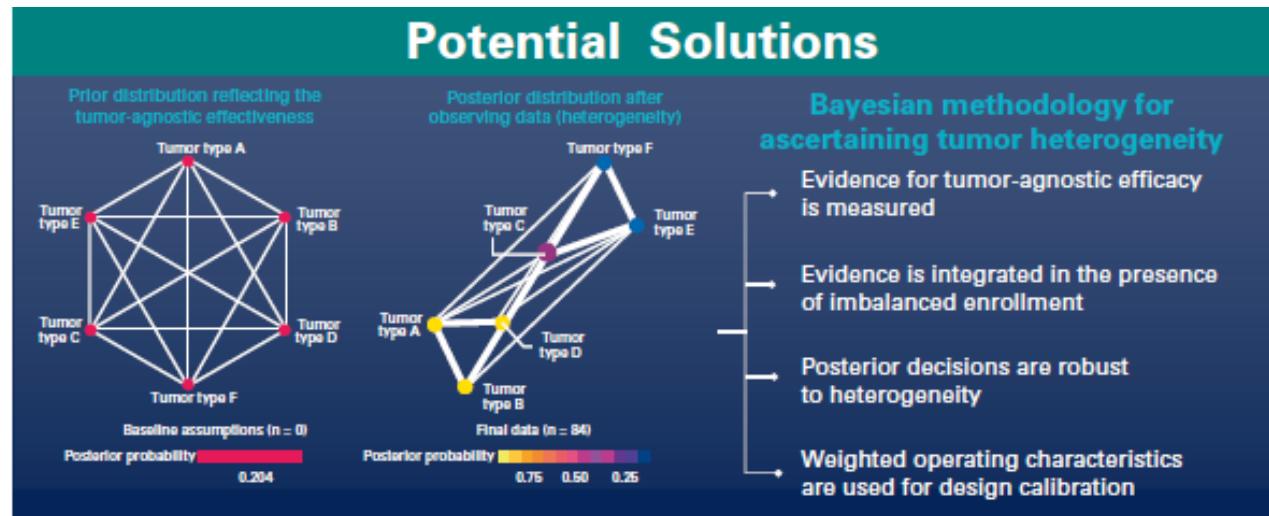
## Basket Designs: Statistical Considerations for Oncology Trials

Alexander M. Kaizer, PhD<sup>1</sup>; Joseph S. Koopmeiners, PhD<sup>2</sup>; Michael J. Kane, PhD<sup>3</sup>; Satrajit Roychoudhury, PhD<sup>4</sup>; David S. Hong, MD<sup>5</sup>; and Brian P. Hobbs, PhD<sup>6</sup>

DEVELOPMENTAL THERAPEUTICS—MOLECULARLY TARGETED AGENTS AND TUMOR BIOLOGY

## Moving Beyond 3+3: The Future of Clinical Trial Design

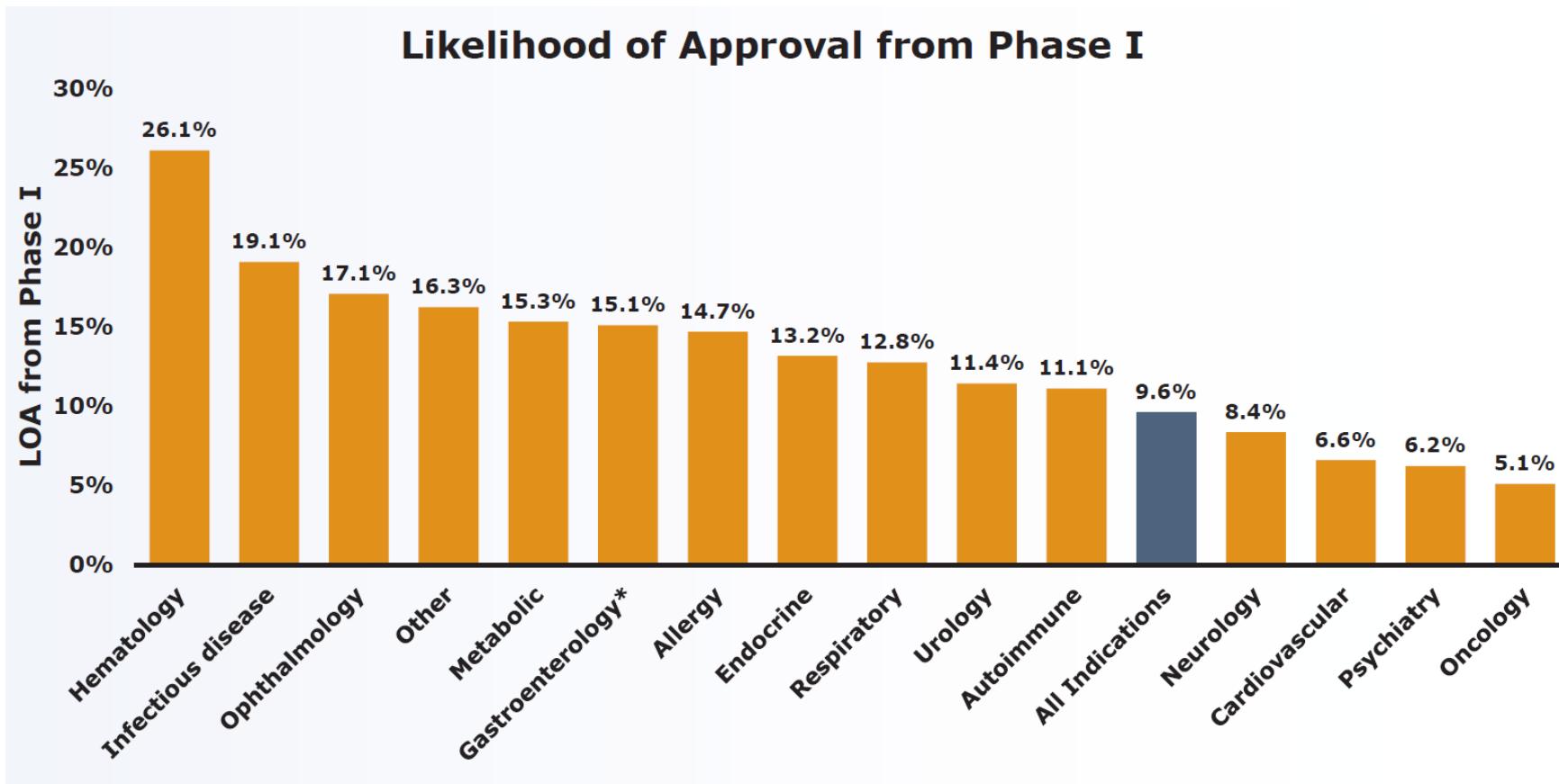
Razelle Kurzrock, MD<sup>1</sup>; Chia-Chi Lin, MD, PhD<sup>2</sup>; Tsung-Che Wu, MD<sup>2</sup>; Brian P. Hobbs, PhD<sup>3</sup>; Roberto Carmagnani Pestana, MD<sup>4</sup>; and David S. Hong, MD<sup>5</sup>



Basket Trials: Review of Current Practice and Innovations for Future Trials  
Hobbs et al. (2022)

Journal of Clinical Oncology®  
An American Society of Clinical Oncology Journal  
Copyright © 2022 by The Journal of Clinical Oncology

# Oncology Low Regulatory Approval Rate



# Sequencing Therapy for Genetically Defined Subgroups of Non-Small Cell Lung Cancer

Helena A. Yu, MD, David Planchard, MD, PhD, and Christine M. Lovly, MD, PhD

A

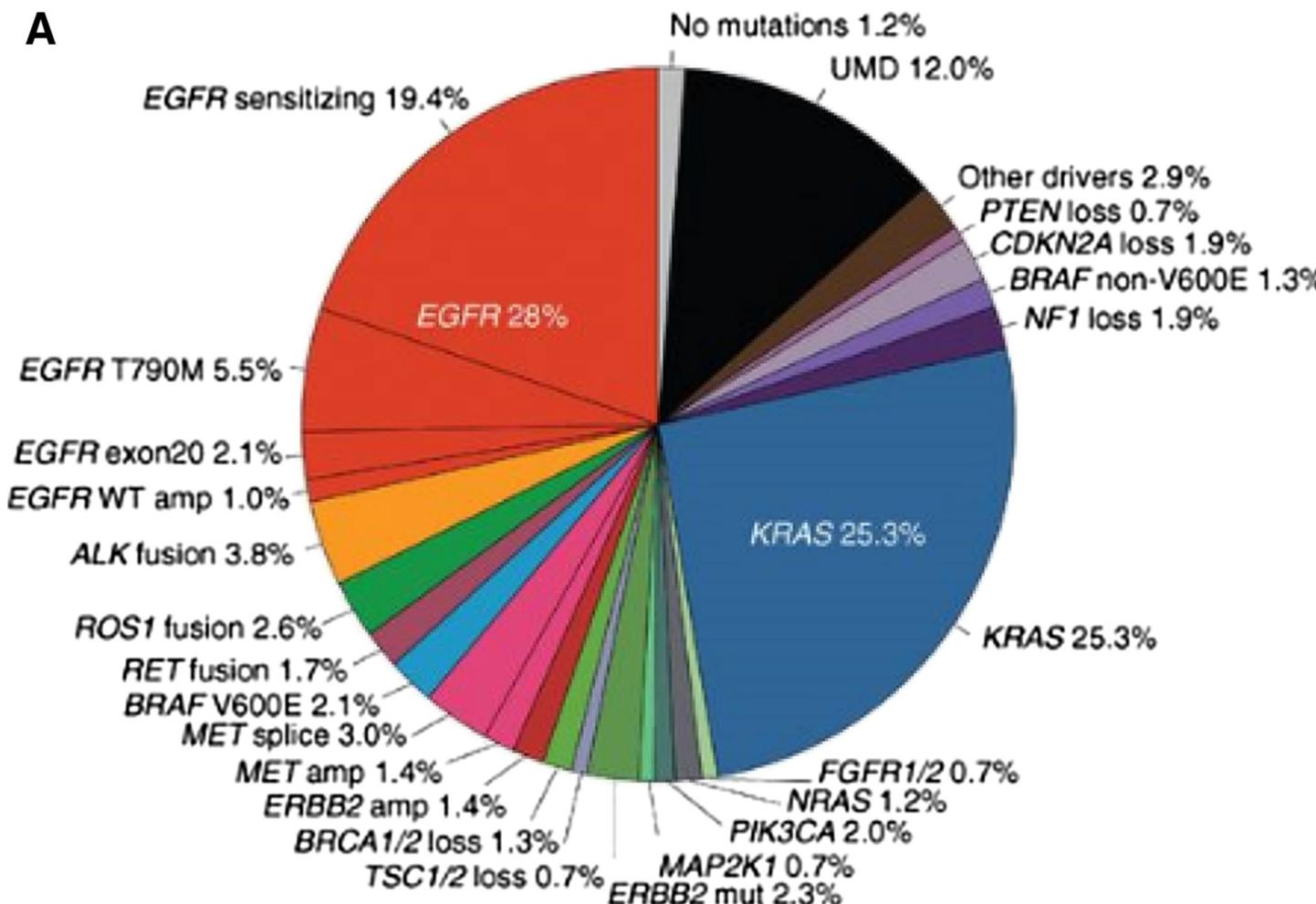


FIGURE 1.

Molecular Cohorts of Lung Cancer

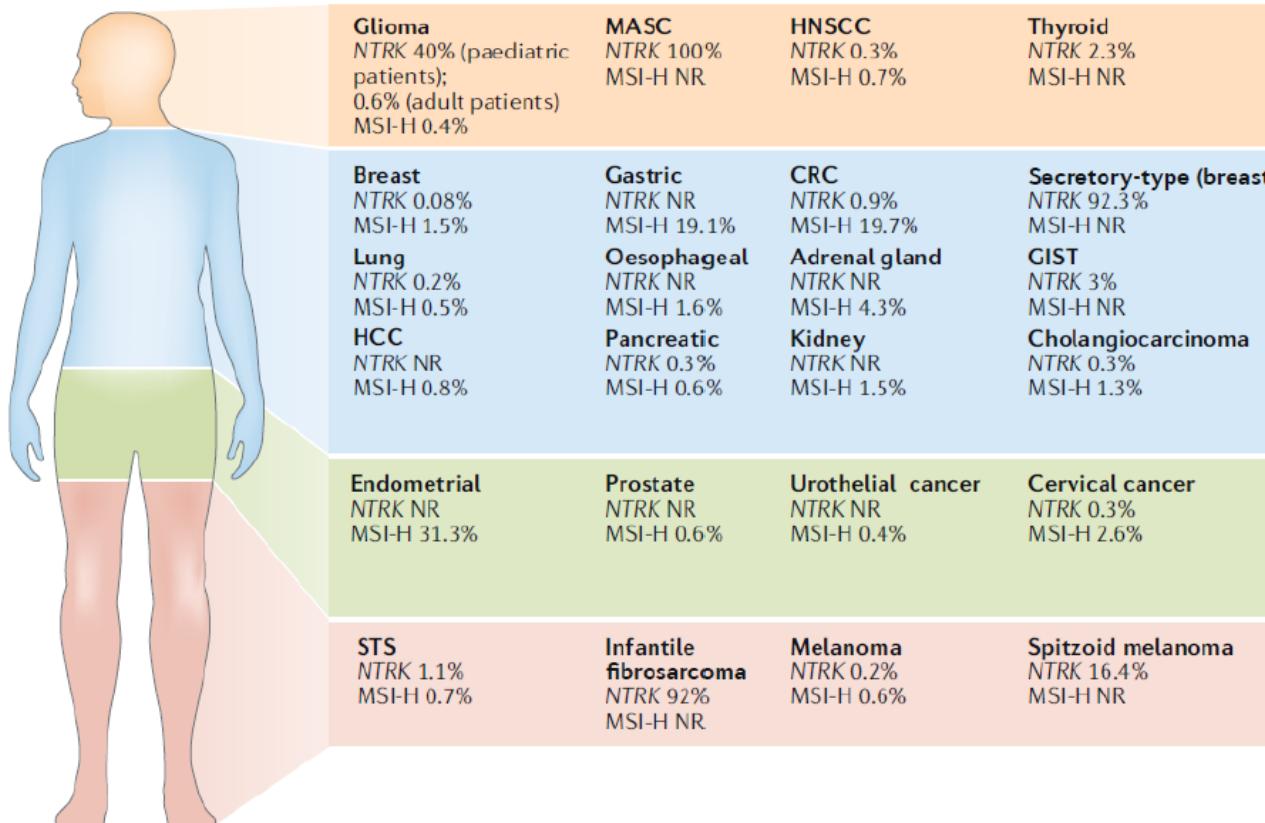
(A) In a large academic center, these are the actionable mutations prospectively identified on a next-generation sequencing mutation platform (430 genes) over a set time period.<sup>103</sup> (B) In a national molecular testing effort, these are the actionable mutations (six-gene panel) prospectively identified over a set time period.<sup>1</sup>

Abbreviation: WT, wild-type.

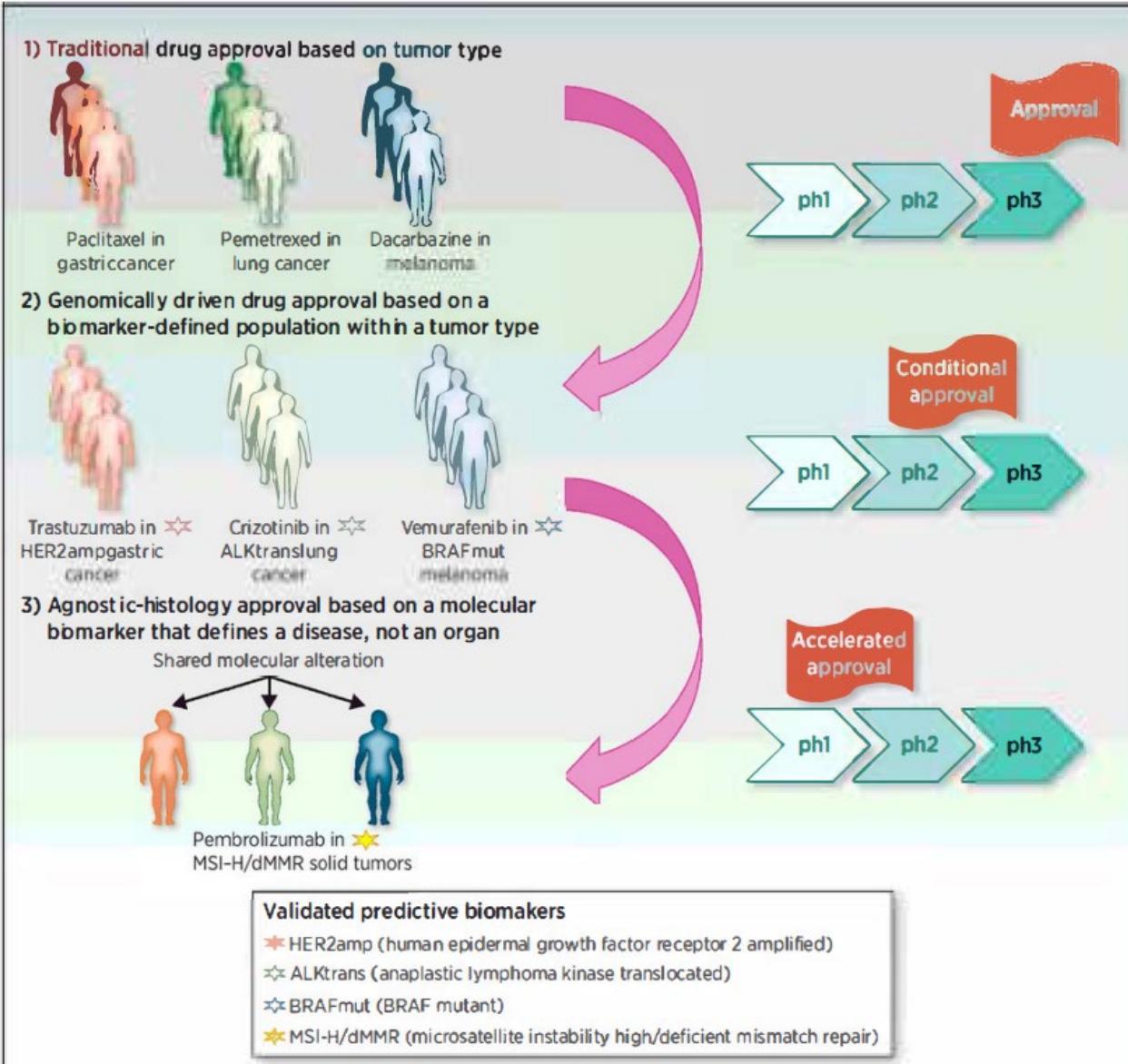
# Histology-agnostic drug development – considering issues beyond the tissue

Roberto Carmagnani Pestana, Shiraj Sen, Brian P. Hobbs & David S. Hong 

*Nature Reviews Clinical Oncology* **17**, 555–568(2020)



**Fig. 2 | Prevalence of specific alterations for which histology-agnostic drugs are approved across tumour types**<sup>33,57,125,153–156</sup>. NTRK gene fusions and microsatellite instability-high (MSI-H) or mismatch-repair deficiency (dMMR) phenotypes are present across multiple tumour types. Knowledge of the prevalence of these features in each tumour



## Regulatory Policy has evolved

Drugs designated as **Breakthrough therapies**, may be approved **without a control group** using **interim results of ongoing studies**, or with **enrichment strategies** devised to select patients who are most likely to benefit

# Withdrawn from Accelerated Approval

Drug Name	Accelerated Approval Date	Withdrawal Date
Ukoniq (umbralisib)	2/5/2021	5/31/2022
Ukoniq (umbralisib)	2/5/2021	5/31/2022
Keytruda (pembrolizumab)	6/17/2019	3/30/2021
Tecentriq (atezolizumab)	3/8/2019	10/6/2021
Copiktra (duvelisib)	9/24/2018	12/17/2021
Opdivo (nivolumab)	8/16/2018	12/29/2020
Keytruda (pembrolizumab)	9/22/2017	2/4/2022
Opdivo (nivolumab)	9/22/2017	7/23/2021
Imfinzi (durvalumab)	5/1/2017	2/19/2021
Tecentriq (atezolizumab)	4/17/2017	12/2/2022
Lartruvo (olaratumab)	10/19/2016	2/25/2020
Tecentriq (atezolizumab)	5/18/2016	4/13/2021
Farydak (panobinostat)	2/23/2015	3/24/2022
Zydelig (idelalisib)	7/23/2014	2/18/2022
Marqibo (vincristine sulfate liposomal)	8/9/2012	5/2/2022
Istodax (romidepsin)	6/16/2011	7/30/2021
Oforta (fludarabine phosphate)	12/18/2008	12/31/2011
Avastin (bevacizumab)	2/22/2008	11/18/2011
Bexxar (tositumomab and iodine i 131 tositumomab)	12/22/2004	10/23/2013
Iressa (gefitinib)	5/5/2003	4/25/2012
Mylotarg (gemtuzumab ozogamicin)	5/17/2000	11/28/2011
Celebrex (celecoxib)	12/23/1999	6/8/2012

# FDA oncology chief aims to open up accelerated approval for earlier cancer treatment under 'Project FrontRunner'

By Angus Liu • Apr 6, 2022 11:15am

U.S. FDA

Richard Pazdur

Oncology drugs

accelerated approval



## Cancer And Accelerated Approval: FDA To Crack Down On Single-Arm Trials, Refractory Disease Focus

10 Jun 2022 | NEWS

### Executive Summary

US FDA cancer chief Rick Pazdur plans to send industry to 'rehab' with Project Frontrunner, which will push for development of cancer drugs in **randomized controlled trials in earlier disease**. Goal is to **reduce time of uncertainty between accelerated approval and confirmatory evidence**.

# Complex Innovative Trial Designs

Center for Biologics Evaluation & Research  
Center for Drug Evaluation & Research

## Innovative Characteristics:

FDA considers the following trial design features to be innovative, making it appropriate to review the design under the Complex Innovative Trial Design (CID) pilot meeting program:

- Use of external control data
- Use of a commensurate prior for borrowing data
- Use of a Bayesian parametric model as the primary analysis of a secondary endpoint

## New! CID Pilot Program Trial Design Case Studies

The description of each CID Pilot Meeting Program case study focuses on the single clinical trial design that was the focus of the Pilot Program submission. The description does not discuss other potentially important aspects of the development program for the respective drug or biologic, such as any plans to conduct additional adequate and well-controlled trial(s) and/or to obtain confirmatory evidence to help establish substantial evidence of effectiveness. Please refer to draft guidance *Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products* (December 2019).

- [Master Protocol Case Study](#)
- [Lupus Case Study](#)
- [DLBCL Case Study](#)

# FDA Takes Important Steps to Increase Racial and Ethnic Diversity in Clinical Trials

*Agency's Focus on Inclusion in Trials for All Medical Products Aligns with Biden Administration's Cancer Moonshot Goal of Addressing Inequities and Beyond*



For Immediate Release: April 13, 2022

Español

Today, the U.S. Food and Drug Administration issued a new draft guidance to industry for developing plans to enroll more participants from underrepresented racial and ethnic populations in the U.S. into clinical trials – expanding on the agency's [previous guidances](#) for industry to improve clinical trial diversity.

**"The U.S. population has become increasingly diverse, and ensuring meaningful representation of racial and ethnic minorities in clinical trials for regulated medical products is fundamental to public health," said FDA Commissioner Robert M. Califf, M.D. "Going forward, achieving greater diversity will be a key focus throughout the FDA to facilitate the development of better treatments and better ways to fight diseases that often disproportionately impact diverse communities. This guidance also further demonstrates how we support the Administration's Cancer Moonshot goal of addressing inequities in cancer care, helping to ensure that every community in America has access to cutting-edge cancer diagnostics, therapeutics and clinical trials."**

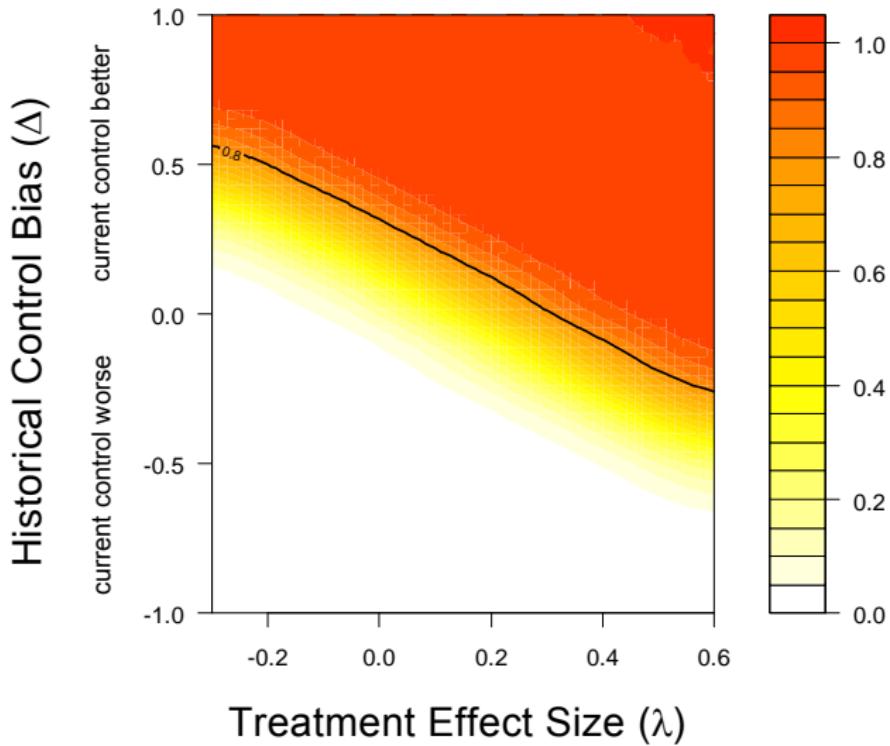
# Background

## Cohort bias

- ▶ unforeseen differences in time, place, and study population
- ▶ variability among the investigators, patient cohorts, and clinical sites
- ▶ inconsistencies in patient supervision or evaluation of therapeutic effects
- ▶ cumulative effects of numerous, small discrepancies
- ▶ diminish comparability among groups of patients risking *biased* treatment comparisons

Not discernable *a priori*

# Power Surface Assuming Exchangeable Data

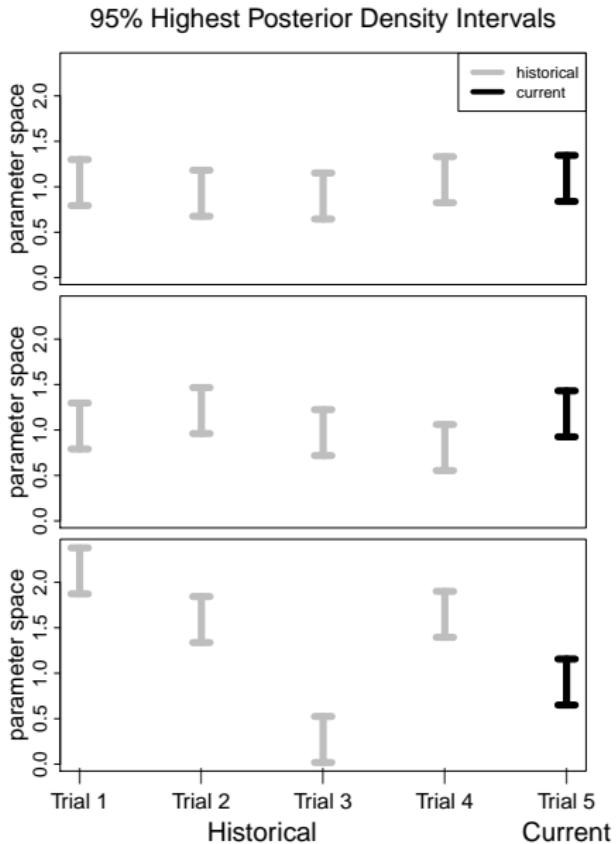


# Pocock's acceptability criteria for historical data

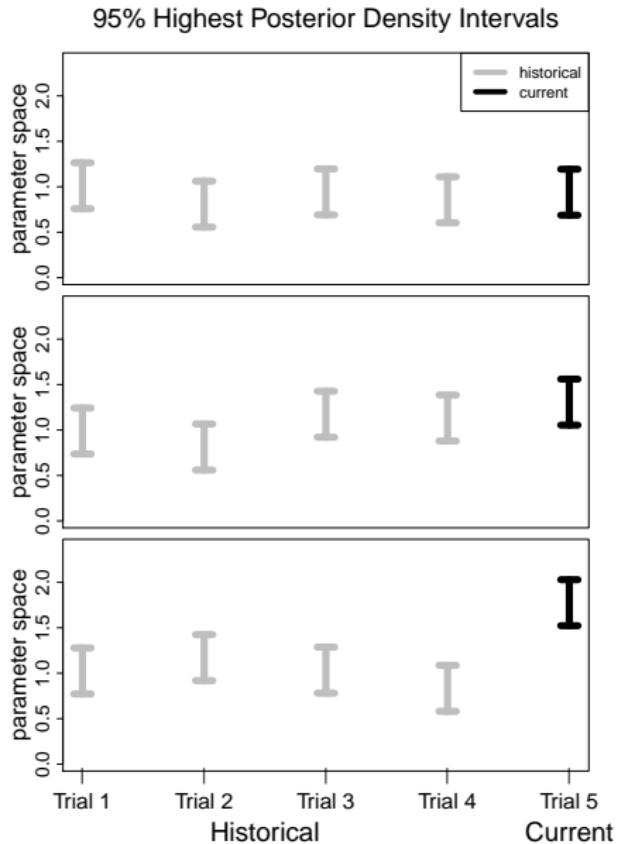
Pocock (1976) *Biometrics*

- ▶ "...presence of *acceptable* historical data cannot be ignored in the full comparative evaluation of a new treatment."
- ▶ proposed the use of both concurrent randomized and historical controls comparative evaluations of novel treatments
- ▶ defined six acceptability criteria
  1. identical therapies
  2. identical requirements for patient eligibility
  3. uniform evaluation of response
  4. comparable patient characteristics
  5. same organization and clinical investigators
  6. no other indications leading one to expect differing results,
    - ▶ i.e. enthusiasm of investigators (patient accrual rates)
- ▶ attempt to eliminate discernable sources of bias
- ▶ yield a putatively *exchangeable* subset of historical trials

## Disperse historical data



## Biased historical data



# Methods of inference

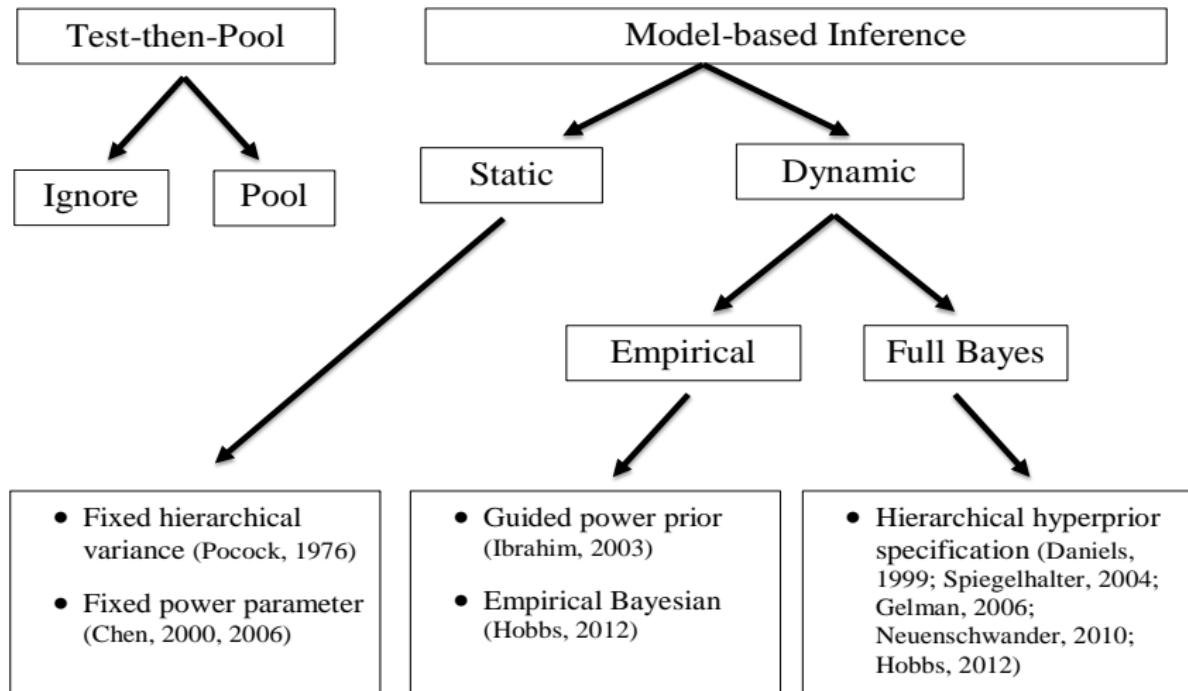
## Static borrowing

- ▶ Bayesian analysis using priors with pre-determined amounts of borrowing
- ▶ prior effective sample size is fixed *a priori* in the absence of the data
- ▶ facilitates data independent partial pooling
- ▶ includes Pocock's proposed approach

## Dynamic borrowing

- ▶ inference whereby the extent of borrowing is not imposed *a priori*, but informed by the data
- ▶ variable effective sample size for the posterior inference (Hobbs et al., 2013)
- ▶ hierarchical modeling facilitates data dependent partial pooling
- ▶ attenuates borrowing in the presence of evidence for trial effects
- ▶ leads to robust estimation with desirable bias/variance trade-offs
- ▶ facilitates adaptive allocation of new patients (Hobbs et al., 2013)

## Schematic of approaches for using historical data



# Model-based inference

## Historical control data

- H historical studies
- $\mathbf{y}_{0j}$  = vector of  $n_{0j}$  responses associated with the current control observed in the  $j^{\text{th}}$  historical study

$$\mathbf{y}_{0j} = \boldsymbol{\mu}_{0j} + \boldsymbol{\epsilon}_{0j}, \quad j = 1, \dots, H$$

$$\boldsymbol{\epsilon}_{0j} \stackrel{\text{iid}}{\sim} \text{Normal}(0, \sigma_{0j}^2)$$

## Current data

- novel treatment compared to the previously studied control
- $y_i$  = response observed in the  $i^{\text{th}}$

$$y_i = \mu + d_i \xi + \epsilon_i, \quad i = 1, \dots, n$$

$$\epsilon_i \sim \text{Normal}(0, \sigma^2)$$

- $d_i$  = indicator for novel treatment;  $n_d = \sum_{i=1}^n d_i$

# Model-based inference

## 1. Irrelevance

- ▶ no borrowing of strength (Spiegelhalter et al., 2004)

## 2. Homogeneity

- ▶ exchangeable data
- ▶ precludes heterogeneity (partial pooling)
- ▶ insensitive to between-trial effects that may overwhelm the inference

## 3. Exchangeable

- ▶ exchangeable control effects
- ▶ standard hierarchical model for meta-analysis (Spiegelhalter et al., 2004)

## 4. Power Prior

- ▶ assume identical effects, but discounts the historical information (Chen et al. 2000)

## 5. Pocock

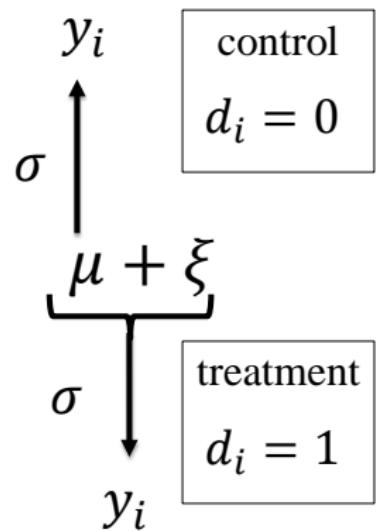
- ▶ considers dispersion of the historical bias (Pocock, 1976)
- ▶ location commensurate prior models are dynamic extension of the Pocock model (Hobbs et al., 2011, 2012, 2013)

## Irrelevance (*no borrowing*) (Spiegelhalter et al., 2004)

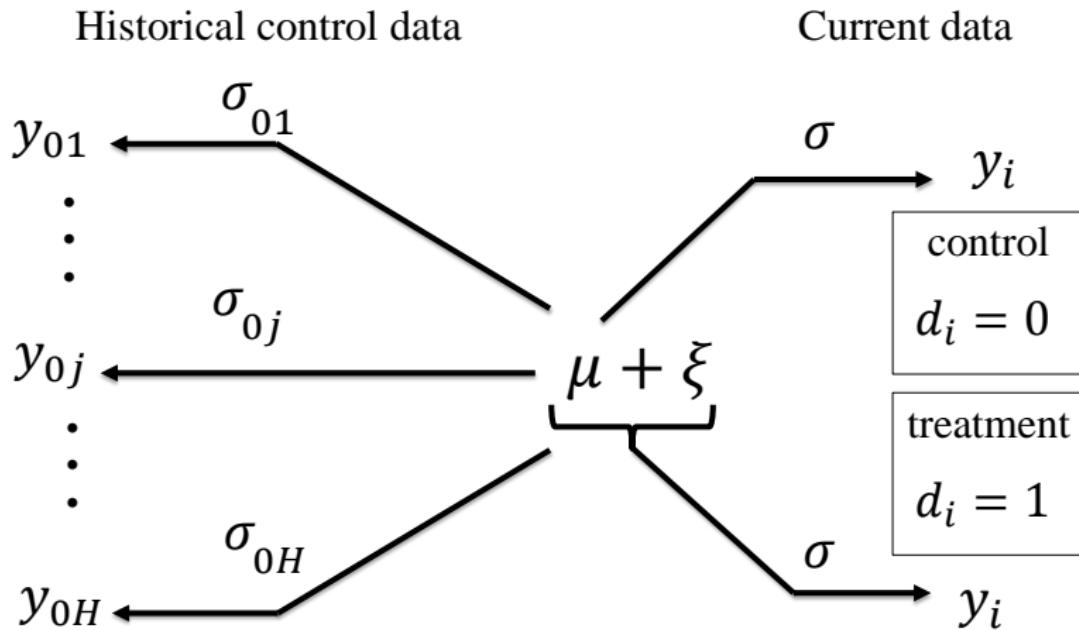
Historical control data

$$\begin{aligned}
 y_{01} &\xleftarrow{\sigma_{01}} \mu_{01} \\
 \vdots &\quad \vdots \\
 \vdots &\quad \vdots \\
 y_{0j} &\xleftarrow{\sigma_{0j}} \mu_{0j} \\
 \vdots &\quad \vdots \\
 \vdots &\quad \vdots \\
 y_{0H} &\xleftarrow{\sigma_{0H}} \mu_{0H}
 \end{aligned}$$

Current data

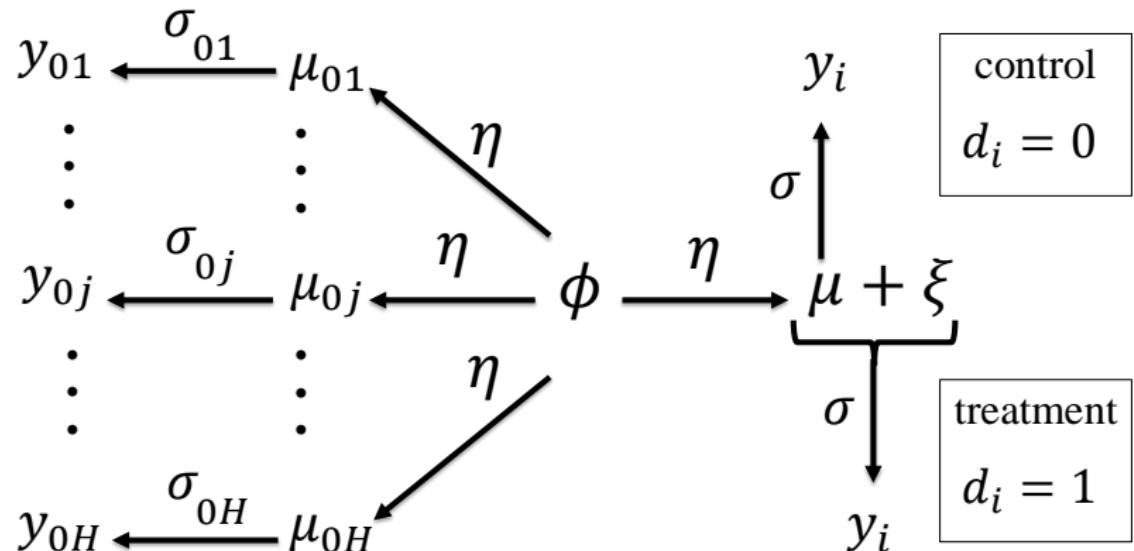


## Homogeneity (*exchangeable* data)



## Exchangeable model (*exchangeable trials*)

Historical control data                      Current data



(Spiegelhalter et al, 2004)

$$B = \sigma^2 / (\sigma^2 + \eta^2)$$

# Exchangeable model (cont'd)

## Prior for current control effect (under flat prior for $\phi$ )

- ▶  $w_{0,h} = (\sigma_{0,h}^2/n_{0,h} + \eta^2)^{-1}$

$$\mu|y_0, \sigma_0^2, \eta^2 \sim N\left(\frac{\sum_{h=1}^H \bar{y}_{0,h}/w_{0,h}^{-1}}{\sum_{h=1}^H w_{0,h}}, \frac{1}{\sum_{h=1}^H w_{0,h}} + \eta^2\right)$$

- ▶ predictive distribution of  $\mu|y_0, \sigma_0^2, \eta^2$  (Neuenschwander, 2010)

## Posterior formulation

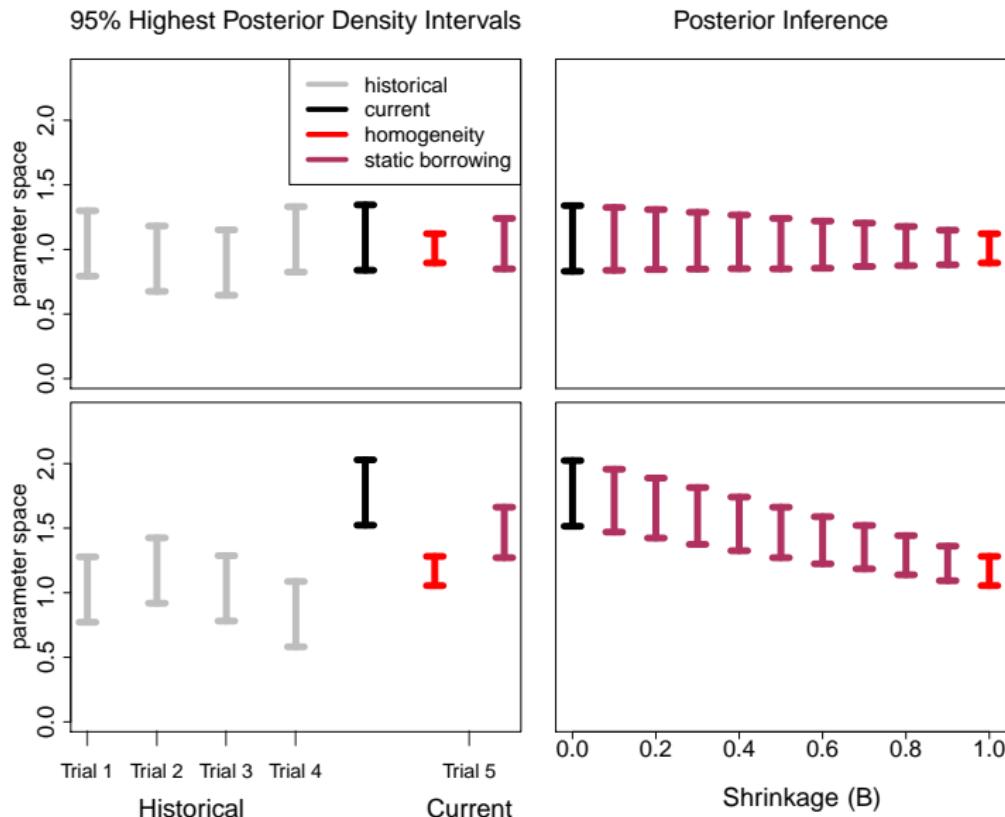
$$\mu|y, y_0, \xi, \sigma^2, \sigma_0^2, \eta^2 \sim N\left(B\phi + (1 - B)(y - \xi), (1 - B)\sigma^2\right), \quad 0 \leq B \leq 1$$

- ▶ Shrinkage parameter:  $B = \sigma^2/(\sigma^2 + \eta^2)$

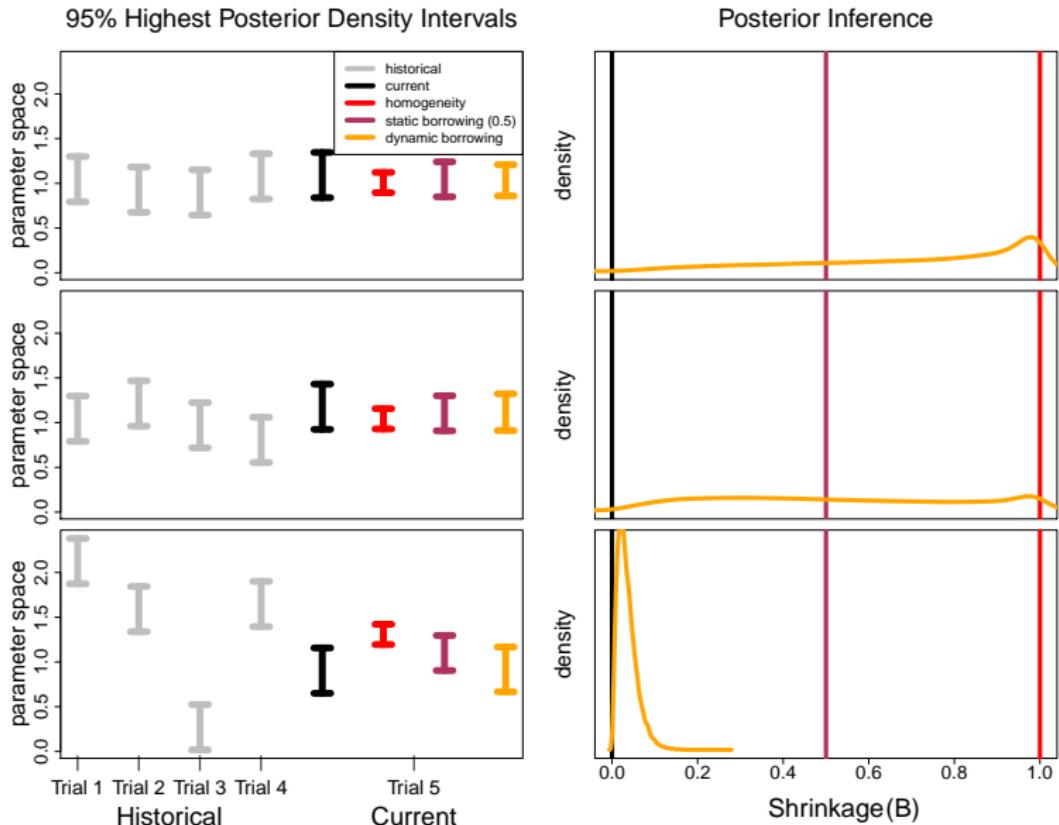
## Methods of inference

- ▶ static inference proceeds with fixed  $\eta^2$
- ▶ dynamic, fully Bayesian inference assumes hyperprior distribution for  $\eta^2$

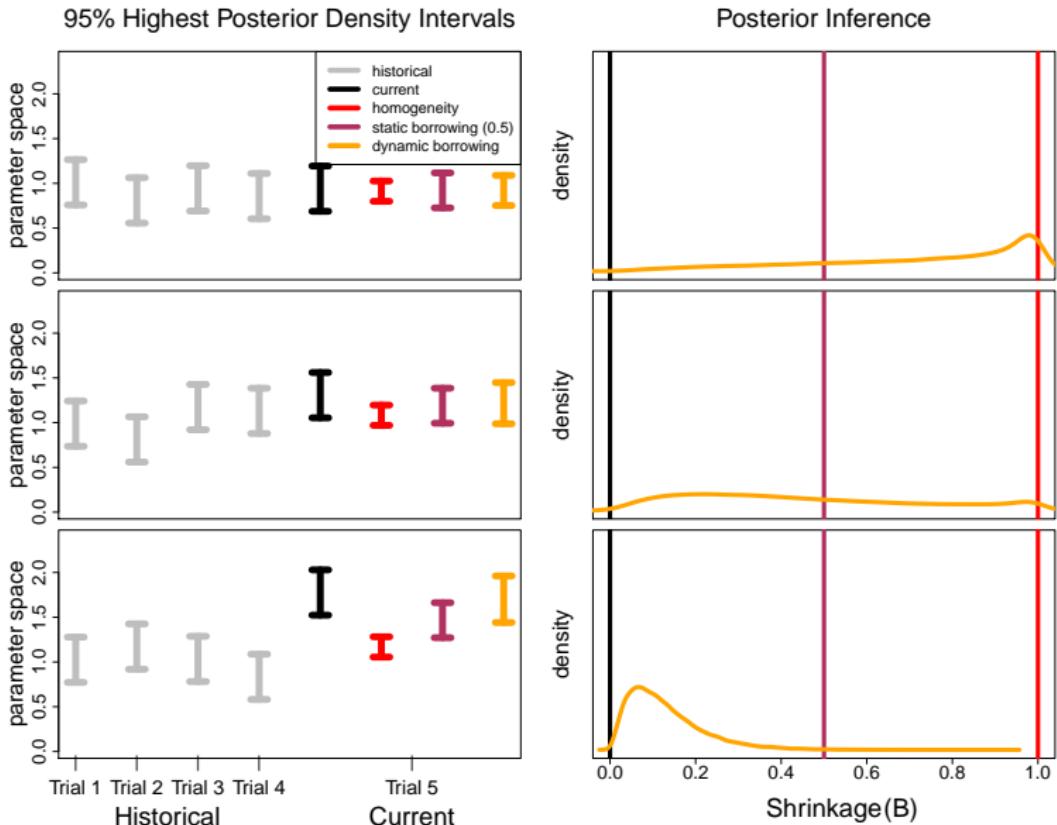
# Static borrowing



# Dynamic borrowing: Disperse historical data



# Dynamic borrowing: Biased historical data



# Dynamic inference

## Hyperprior specification

Strawderman, 1971; Daniels, 1999

- ▶ Uniform(0,1) prior on the shrinkage parameter,  $B$
- ▶ induces vague proper prior for the variance component,  $\eta^2$

Spiegelhalter et al., 2004

- ▶ conditionally conjugate inverse gamma for  $\eta^2$
- ▶ uniform variance ( $\eta^2$ ) or standard deviation ( $\eta$ )

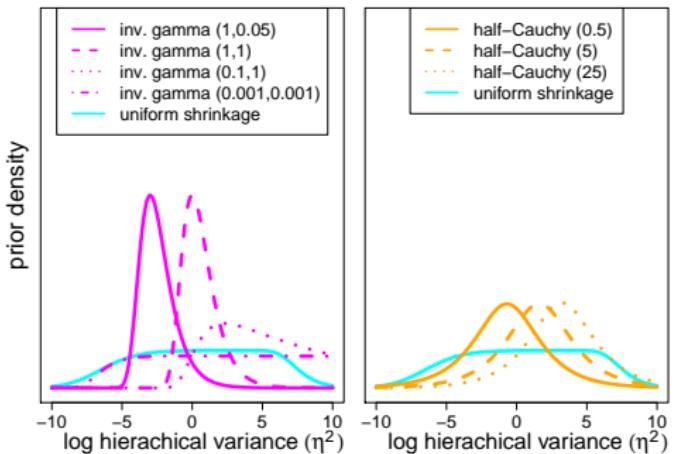
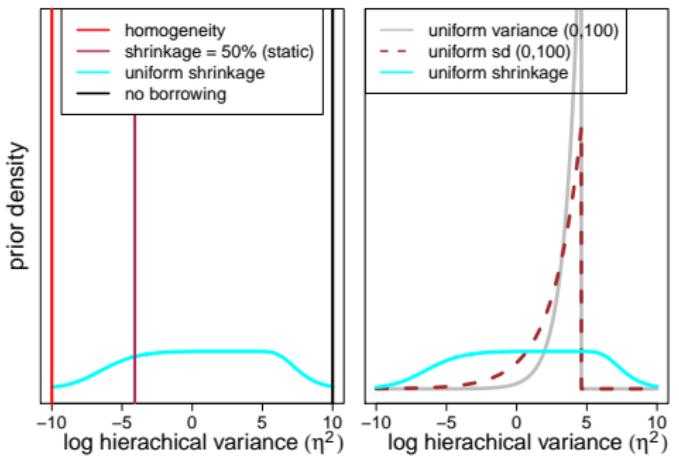
Gelman, 2006

- ▶ uniform prior on  $\eta$  is *improper* for  $H < 3$
- ▶ does not recommend the inverse gamma prior in this context
- ▶ provides alternative, *proper* half-Cauchy prior for  $\eta$ 
  - ▶ a special case of the conditionally-conjugate folded-noncentral-t family

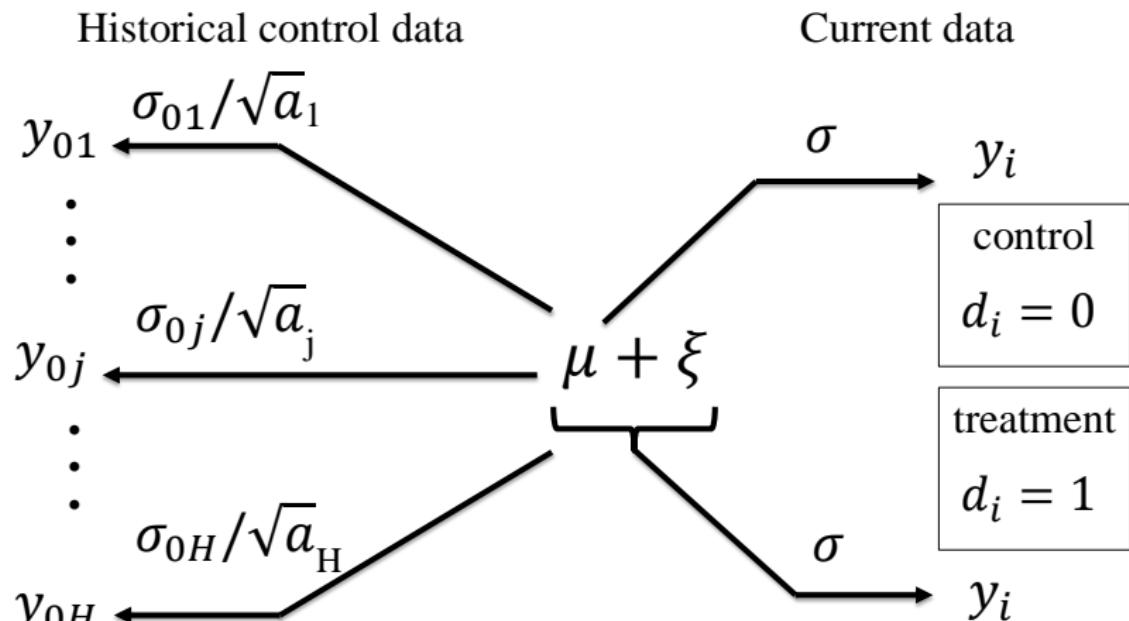
Common “noninformative” and “weakly-informative” hyperpriors

prior	formulation
uniform variance <sup>a</sup>	$p(\eta^2) = U(0, a), a = 100$
<b>inverse gamma</b> <sup>a</sup>	$p(\eta^2) = \Gamma^{-1}(\epsilon, \epsilon), \epsilon = 0.001$
uniform standard deviation <sup>a</sup>	$p(\eta) = U(0, \sqrt{a})$
<b>half-Cauchy</b> <sup>b</sup>	$p(\eta) \propto (\eta^2 + b)^{-1}$
<b>uniform shrinkage</b> <sup>c</sup>	$p(\eta^2) \propto \sigma^2 / \{(\sigma^2 + \eta^2)^2\}$

<sup>a</sup> = Spiegelhalter et al., 2004; <sup>b</sup> = Gelman, 2006; <sup>c</sup> = Strawderman, 1971; Daniels, 1999



## Power Prior (*identical but discounted historical data*)



(Chen 2000; Ibrahim 2003; Hobbs 2011)

$$0 \leq a_j \leq 1, \forall j$$

# Power prior model (cont'd)

## Prior for current control effect

- ▶  $p(\mu|y_0, \sigma_0^2, a_0) \propto L(\mu|y_0)^{a_0} p(\mu)$ ,  $0 \leq a_0 \leq 1$  (Ibrahim et al., 2003)
- ▶ under flat initial prior for  $\mu$ ,  $p(\mu)$

$$\mu|y_0, \sigma_0^2, a_0 \sim N\left(\bar{y}_0, a_0^{-1}(\sigma^2/n_0)\right)$$

## Methods of inference

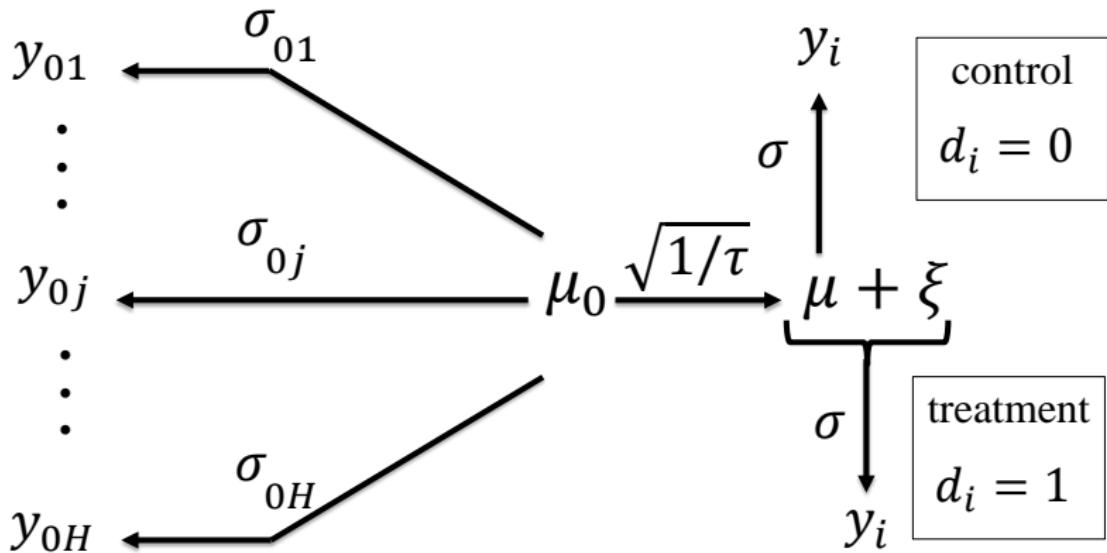
- ▶ *static* inference proceeds under fixed  $a_0$
- ▶ *dynamic, empirical Bayesian* proceeds with inference using a value of  $a_0$  fixed to satisfy pre-specified criterion based upon the observed data
- ▶ Ibrahim et al. (2003) propose *guide value* minimizing penalized likelihood-type criterion

$$G(a_0) = -2 \log(h^*(a_0)) + \frac{\log(n_0)}{a_0}$$

- ▶ where  $h^*(a_0) = \int L(\mu|y)L(\mu|y_0)^{a_0} p(\mu) d\mu$
- ▶ Hobbs et al. (2011) considered an alternative dynamic approach

## Pocock model (exchangeable historical data)

Historical control data                      Current data



(Pocock, 1976; Hobbs et al., 2011, 2012)

*historical bias:  $\Delta = \mu - \mu_0$*

# Pocock model (cont'd)

## Prior for current control effect $\mu$ (under flat prior for $\phi$ )

- ▶  $u_{0,h} = n_{0,h}/\sigma_{0,h}^2$  independent of  $\tau$

$$\mu | \mathbf{y}_0, \boldsymbol{\sigma}_0^2, \tau \sim N \left( \sum_{h=1}^H \frac{\bar{y}_{0,h}/u_{0,h}^{-1}}{\sum_{h=1}^H u_{0,h}}, \frac{1}{\sum_{h=1}^H u_{0,h}} + \frac{1}{\tau} \right)$$

## Posterior formulation

- ▶ ignoring current:  $\mu_0 | \boldsymbol{\sigma}_0^2, \mathbf{y}_0 \sim N(\hat{\mu}_0, v_0)$ , where  $v_0 = (\sum_{h=1}^H u_{0,h})^{-1}$
- ▶ ignoring historical:  $\mu | \boldsymbol{\sigma}^2, \mathbf{y} \sim N(\hat{\mu}, \sigma^2 / (n - \sum_i d_i))$
- ▶ joint inference:

$$\mu | \tau, \boldsymbol{\sigma}_0^2, \sigma^2, \mathbf{y}_0, \mathbf{y} \sim N \left\{ C\hat{\mu}_0 + (1 - C)\hat{\mu}, (v_0 + \frac{1}{\tau})C \right\}$$

- ▶  $C = v/(v + v_0 + \frac{1}{\tau})$  weight associated with  $\hat{\mu}_0$

# Pocock model (cont'd)

## Methods of inference

- ▶ *static* inference proceeds with fixed value of  $\tau$  (Pocock, 1976)
- ▶ *dynamic, empirical Bayesian* fixes  $\tau$  at MMLE (Hobbs et al., 2012)
- ▶ *dynamic, fully Bayesian* assumes prior distribution for  $\tau$  (Hobbs et al., 2011, 2012)

## Estimation of $\tau$

- ▶  $\hat{\Delta} = \hat{\mu} - \hat{\mu}_0$  *estimated historical bias*
- ▶ marginal likelihood as a function of  $\tau$

$$m(\mathbf{y}, \mathbf{y}_0 | \tau) \propto N \left\{ \hat{\Delta} \mid 0, v + v_0 + \frac{1}{\tau} \right\}$$

- ▶  $\hat{\Delta}$ ,  $v$ , and  $v_0$  are *independent of  $\tau$*
- ▶  $\tau$  accounts for extra-dispersion associated with the historical bias

# Pocock Model: Dynamic Inference

## Parametric empirical Bayesian (Hobbs et al., 2012)

- fix  $\tau = 1/\nu$  at MMLE restricted to bounded domain,  $0 < l_\nu < u_\nu$ ,

$$\tau = \left( \arg \max_{\nu \in [l_\nu, u_\nu]} \{m(\mathbf{y}, \mathbf{y}_0 | 1/\nu)\} \right)^{-1}$$

- bounding precludes homogeneity

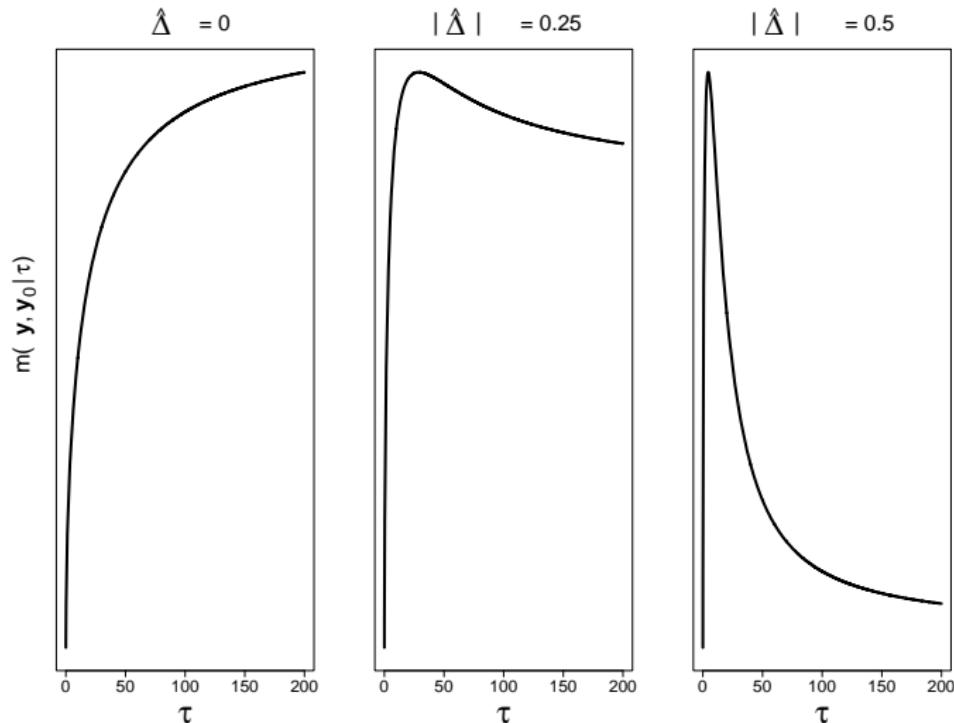
## Fully Bayesian (Hobbs et al., 2012)

- assume spike-and-slab prior for  $\tau$  (Mitchell and Beauchamp, 1988)
- locally uniform between  $0 \leq \mathcal{S}_l < \mathcal{S}_u$
- with additional point mass,  $\mathcal{K} > \mathcal{S}_u$ , such that

$$Pr(\tau > \mathcal{S}_u) = Pr(\tau = \mathcal{K}) = 1 - p_0$$

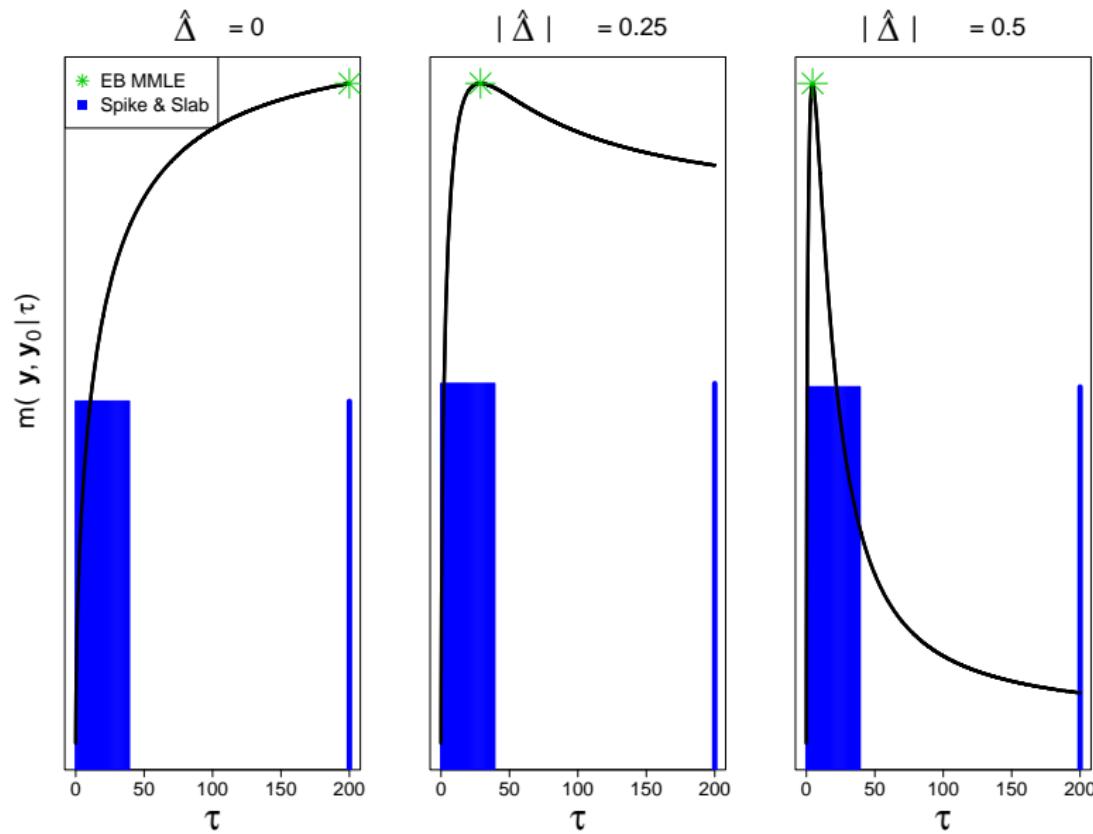
- where  $p_0$  denotes the prior point-mass prob.  $\mathcal{S}_l \leq \tau \leq \mathcal{S}_u$

# Marginal density



Marginal density of  $\mathbf{y}, \mathbf{y}_0 | \tau$  as a function of  $\tau$  for three values of the estimated historical bias  $|\hat{\Delta}|$

# Marginal density and prior



# Effective historical sample size

## Notation:

- ▶  $\mathcal{P}^*(y)$  = precision of  $\mu|y$  under the reference model
- ▶  $\mathcal{P}(y_0, y)$  = precision of  $\mu|y_0, y$  under the joint model
- ▶  $n_c = n - \sum_i di$  number of concurrent controls

## Precision in relation to sample size under reference model:

- ▶  $\mathcal{P}^*(y) = g(n_c) = n_c/\sigma^2$

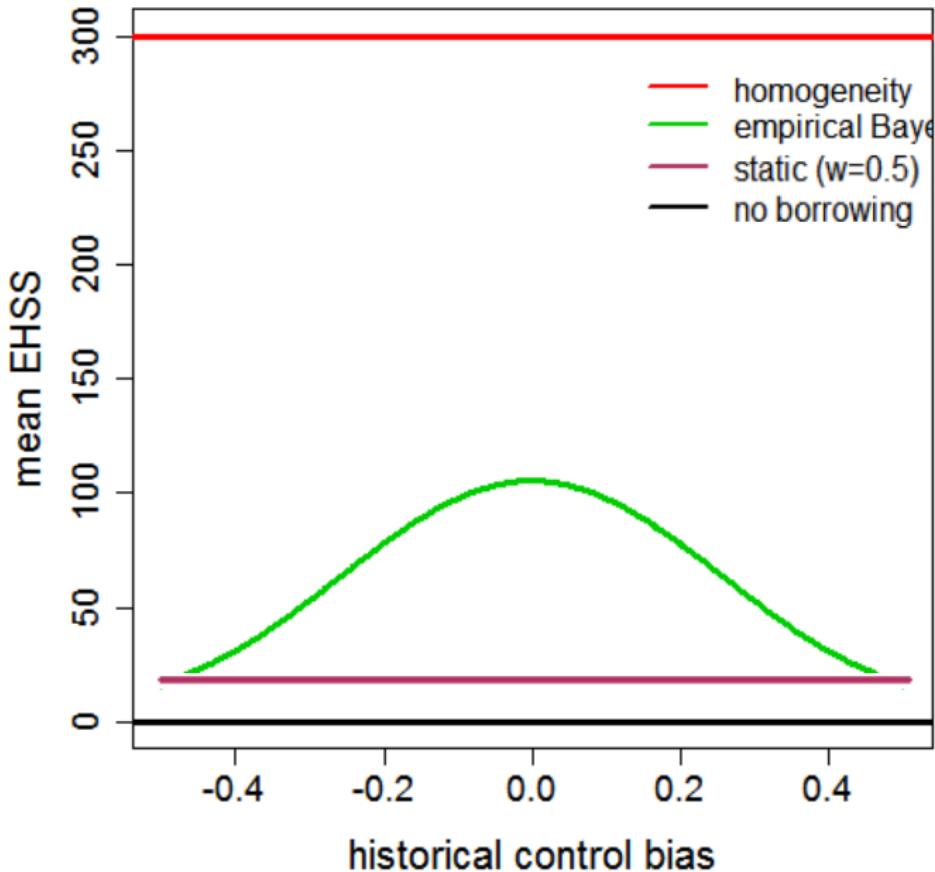
## Effective sample size of joint inference (*a posteriori*):

- ▶  $g^{-1}\{\mathcal{P}(y_0, y)\} = \sigma^2 \mathcal{P}(y_0, y) = n_c + \frac{\sigma^2}{v_0 + \frac{1}{\tau}}$

## Effective *historical* sample size

$$EHSS = g^{-1}\{\mathcal{P}(y_0, y)\} - n_c = \frac{\sigma^2}{v_0 + \frac{1}{\tau}}$$

## Effective historical sample size



# Properties for estimation of novel treatment effect $\xi$

## Notation

- ▶  $\theta^{true}$  = set of fixed, “true” values of the model parameters
- ▶  $\mathcal{D} = (y_0, y)$  denotes data

## Frequentist properties

1. bias:  $E_{\mathcal{D}|\theta^{true}} \{E_{\xi|\mathcal{D}}(\xi)\} - \xi^{true}$
2. risk under squared error loss:  $E_{\mathcal{D}|\theta^{true}} \left[ \{E_{\xi|\mathcal{D}}(\xi) - \xi^{true}\}^2 \right]$
3. coverage of interval estimation of  $\xi$  using 95% HPD interval
4. width of 95% HPD interval for  $\xi$

## Bayesian properties

1. probability allocated to decision rules in the presence of indifference region

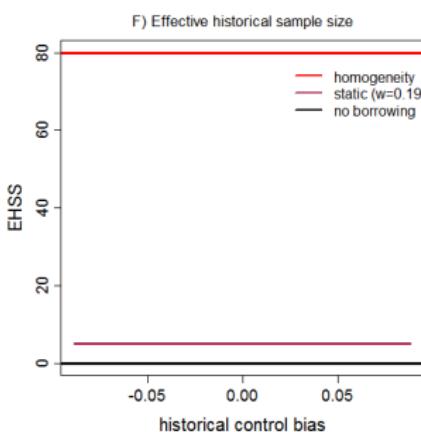
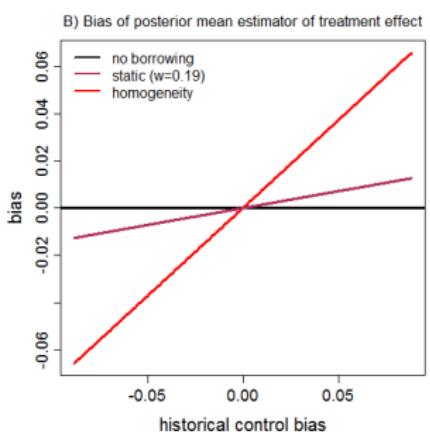
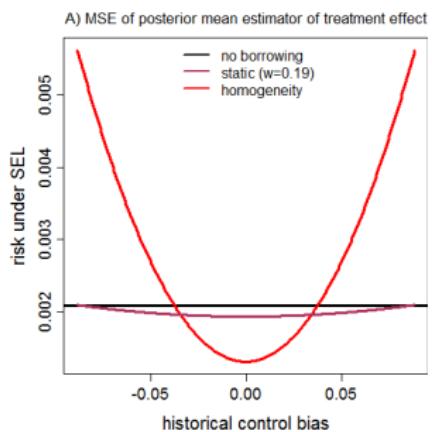
# Characterizing Bias

## Single historical study

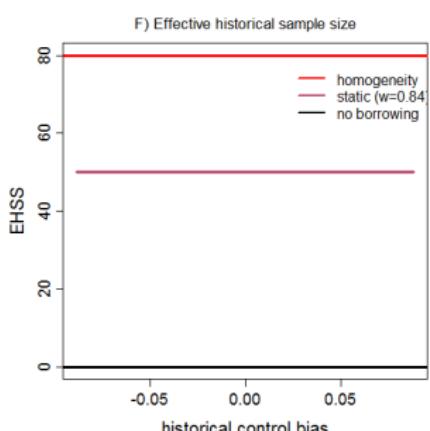
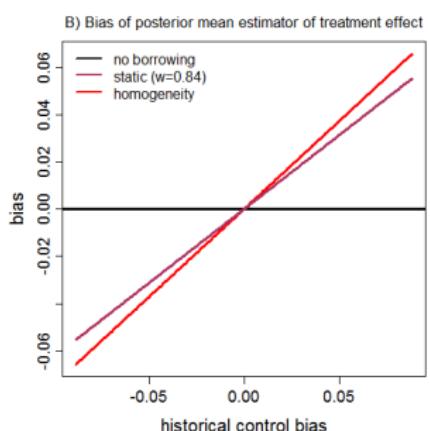
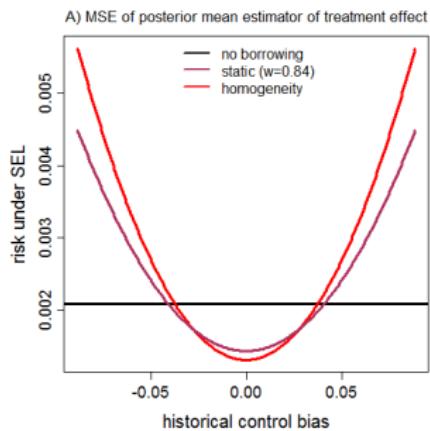
- ▶ consider fixed magnitudes of the inter-trial effect,  $\Delta$ , ( $-\Delta$  = historical bias)

## Compare freq. properties under inference using

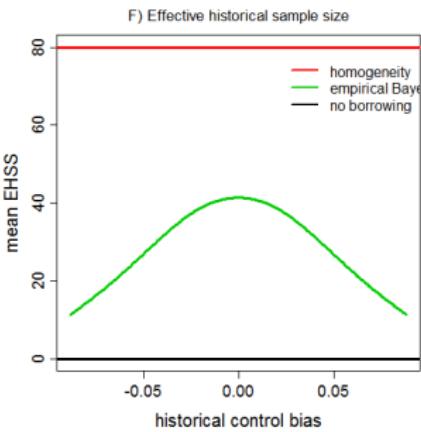
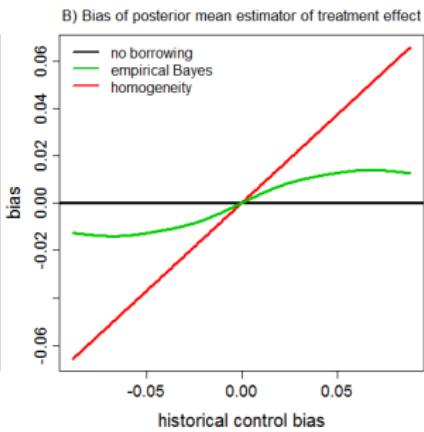
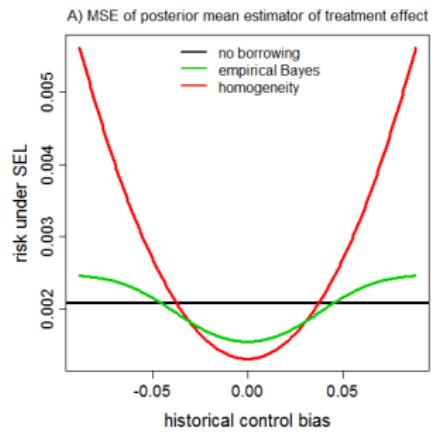
1. Power Prior model
  - ▶ *static* inference with fixed  $a_0$
  - ▶ *dynamic*, empirical “guided” implementation
2. Pocock model
  - ▶ *static* inference with fixed  $\tau$
  - ▶ *dynamic*, empirical and fully Bayesian spike-and-slab implementations for estimating  $\tau$



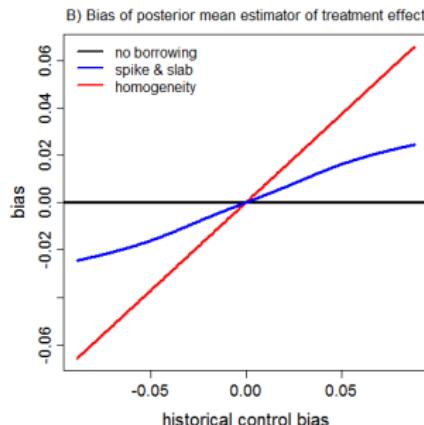
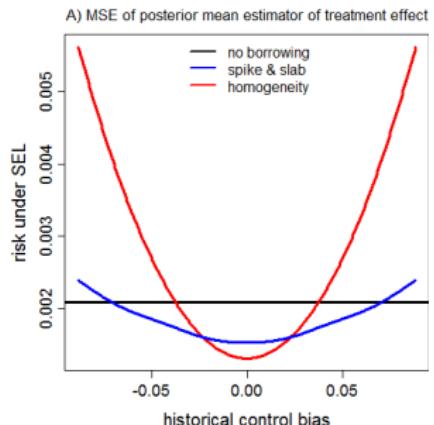
(a)



(b)



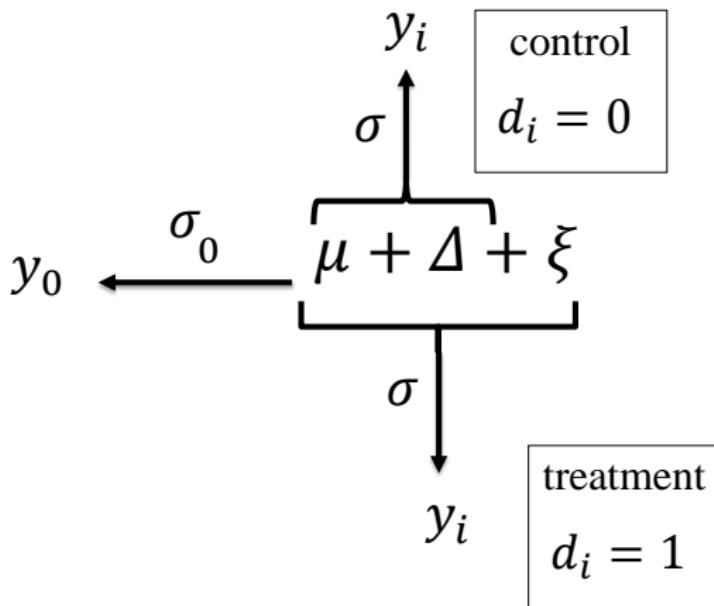
(a)



(b)

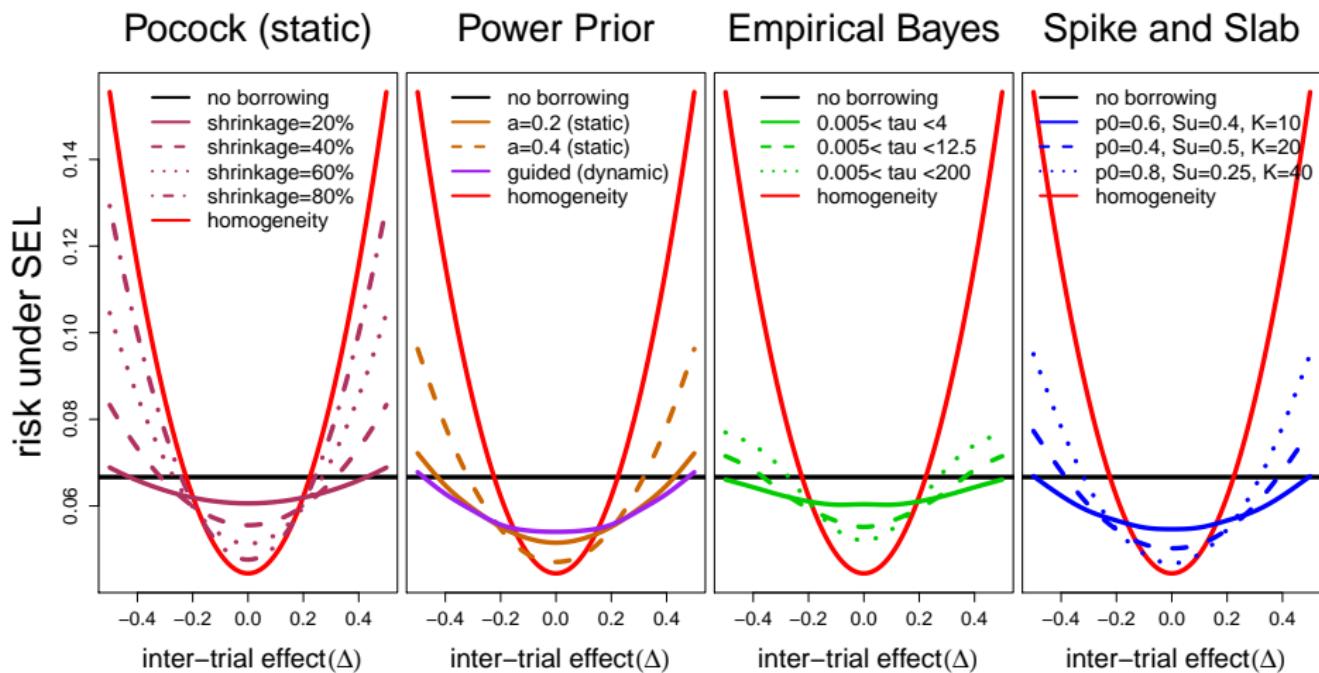
## Bias model (*historical bias in presence of a single historical study*)

Historical control data      Current data

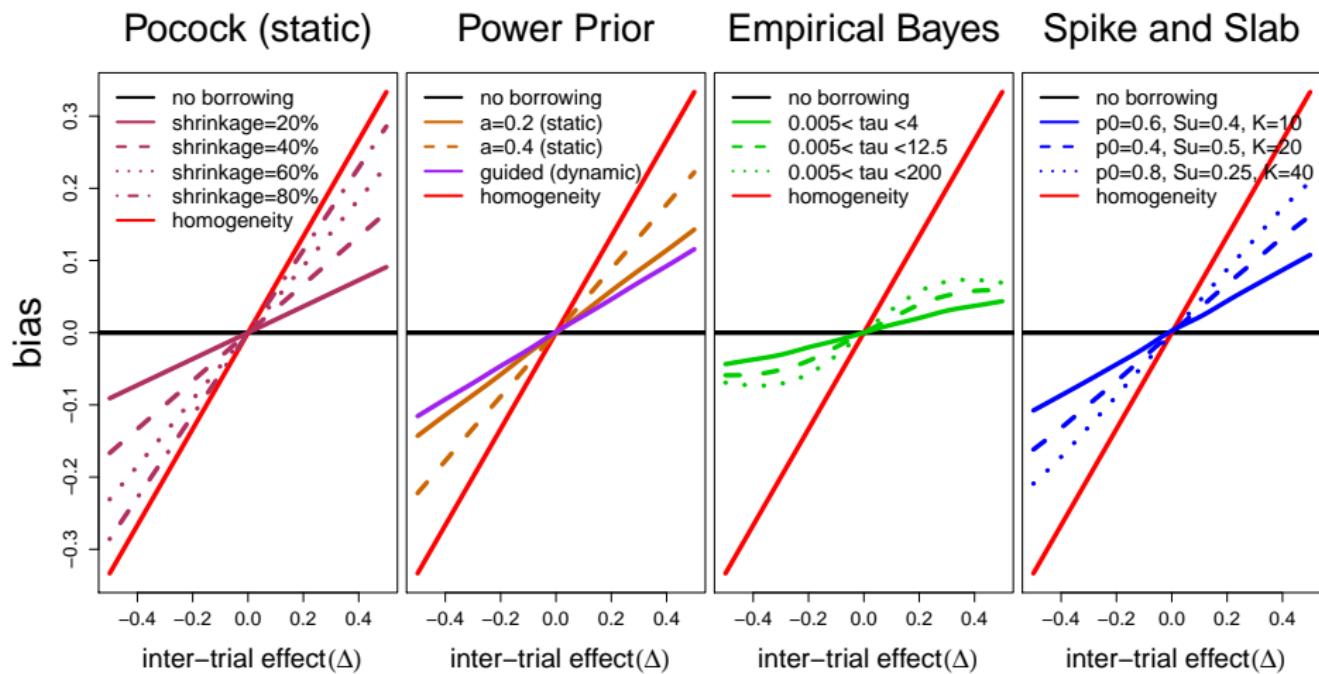


\*used in simulation study of frequentist operating characteristics

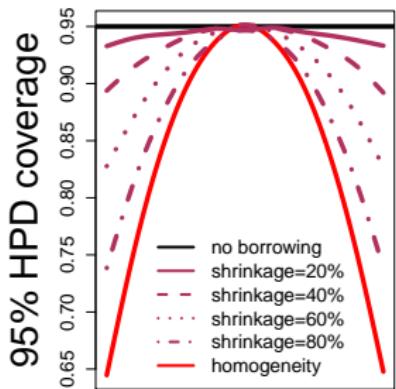
# Comparison of frequentist properties: Risk under SEL



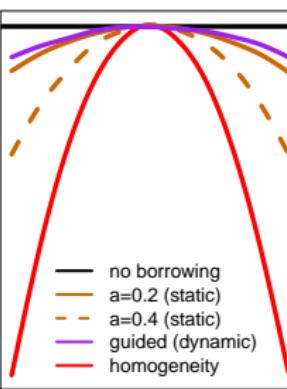
# Comparison of frequentist properties: Bias



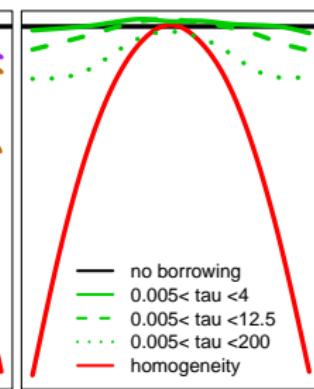
Pocock (static)



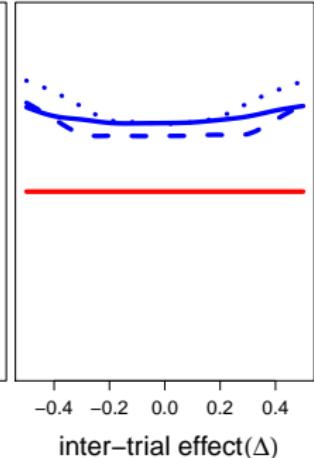
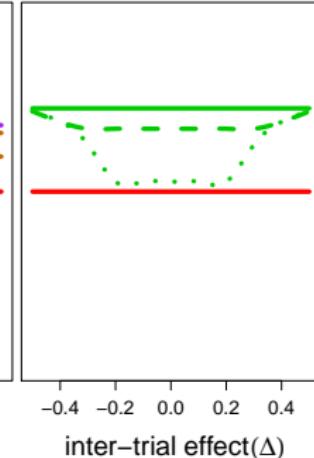
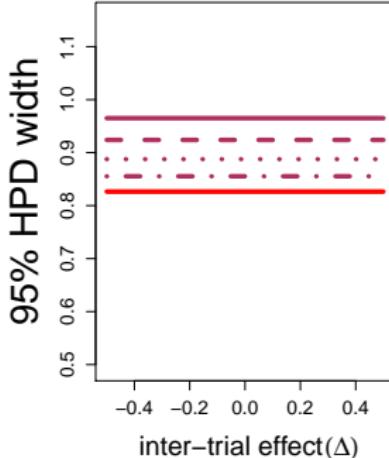
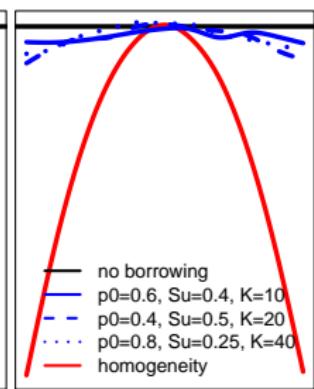
Power Prior



Empirical Bayes



Spike and Slab



# Characterizing heterogeneity

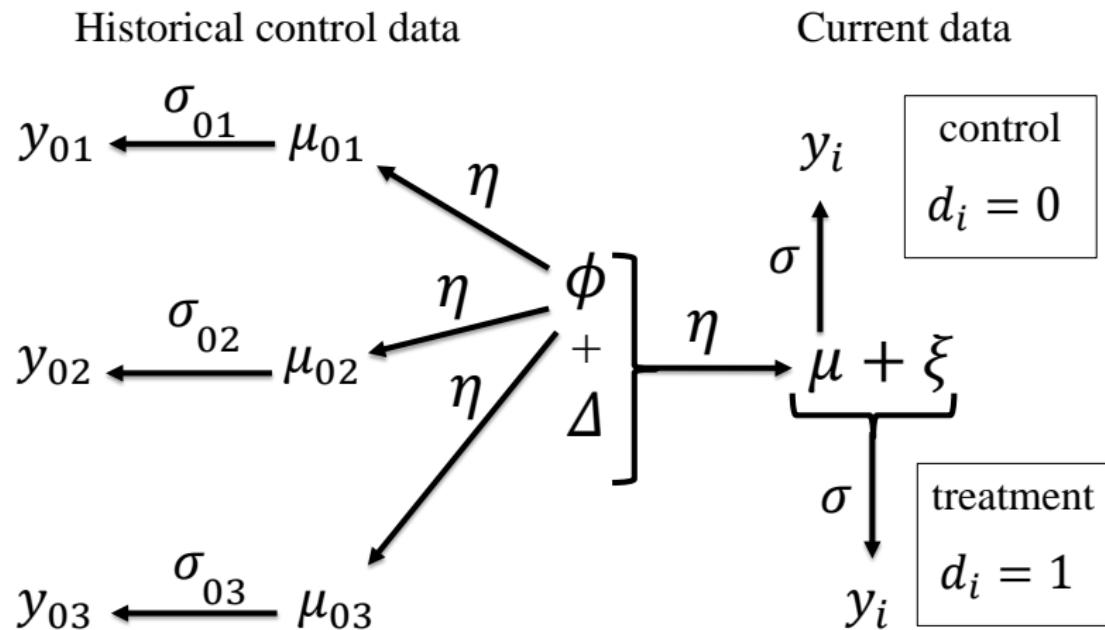
## Multiple historical studies

- ▶ dispersion of trial effects via hierarchical standard deviation,  $\eta$
- ▶ historical trial effects are assumed to vary non-systematically about hierarchical mean,  $\phi$
- ▶ consider fixed magnitudes of the current control effect  $\Delta$

## Compare freq. properties under inference using

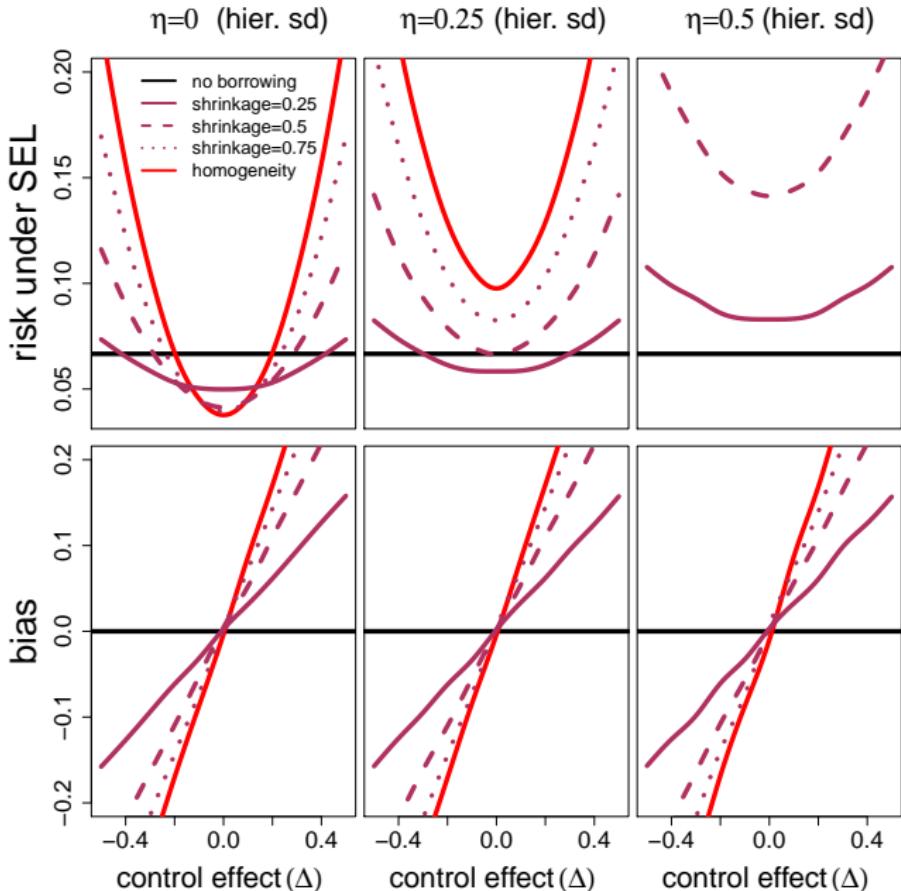
1. Exchangeable model
  - ▶ static inference with fixed shrinkage,  $B$
2. Pocock model
  - ▶ dynamic, empirical Bayesian implementation for estimating  $\tau$

## Dispersion-Bias model (*historical dispersion and current bias*)

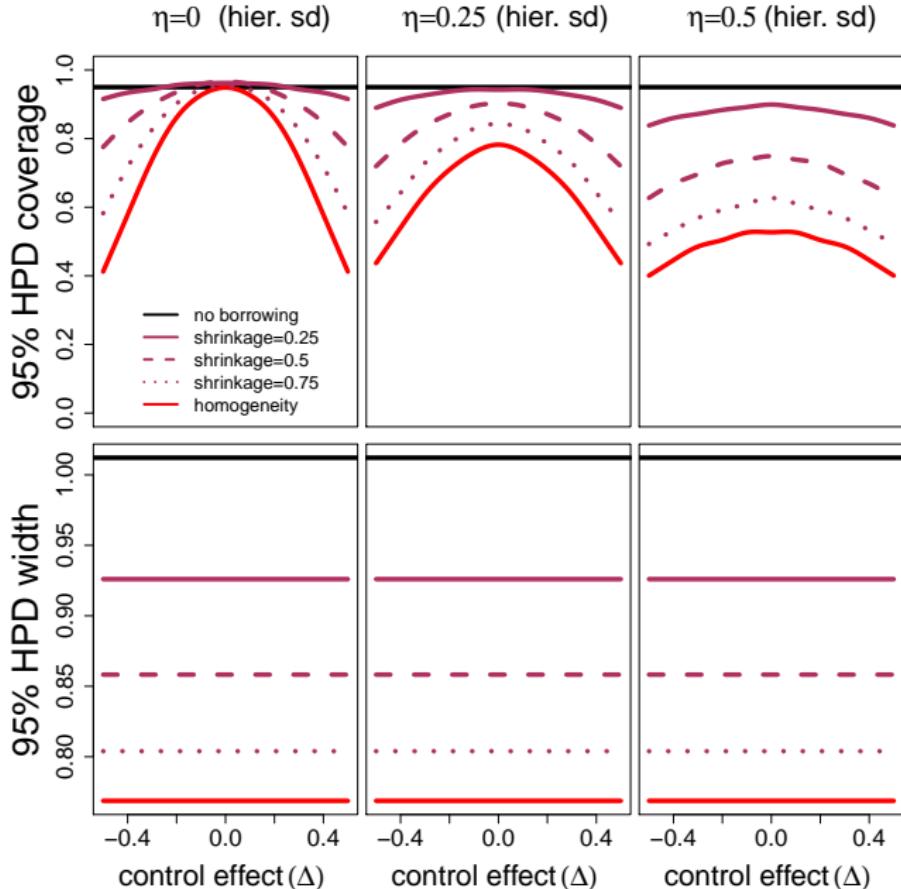


\*used in simulation study of frequentist operating characteristics

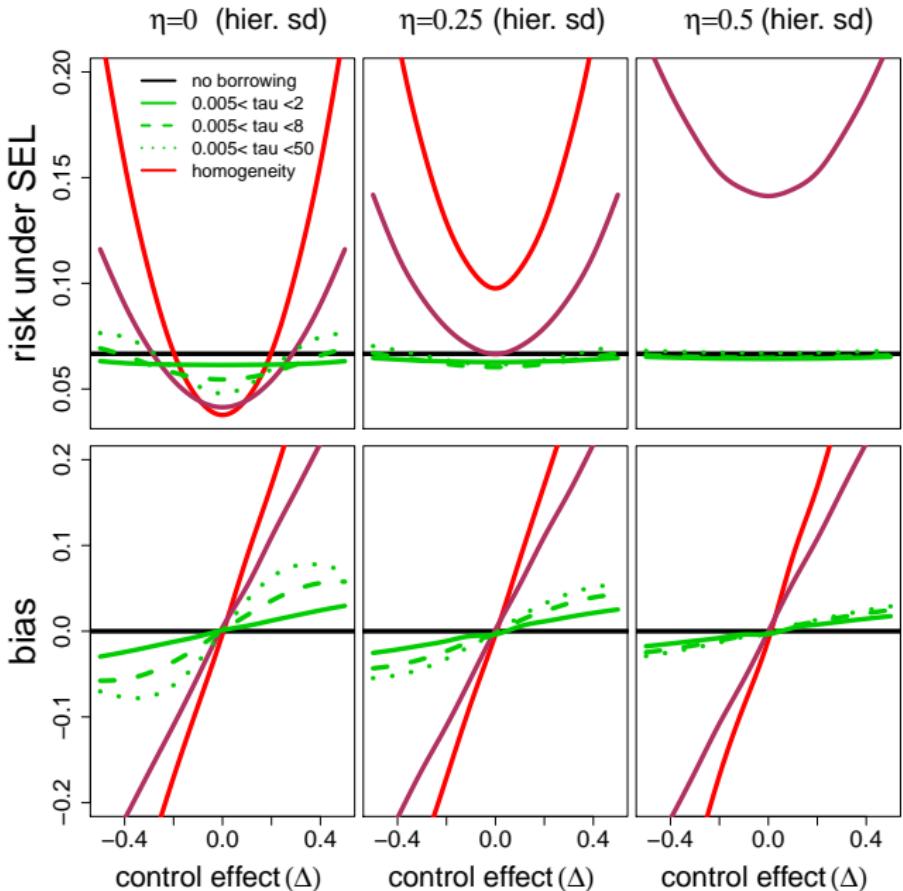
## Exchangeable model: static borrowing (fixed shrinkage)



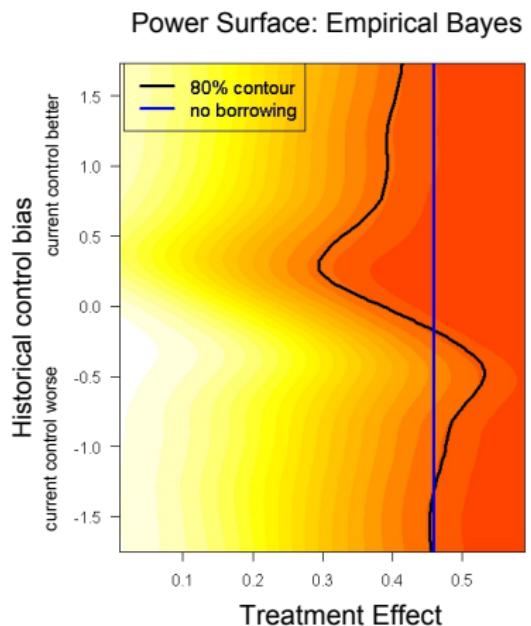
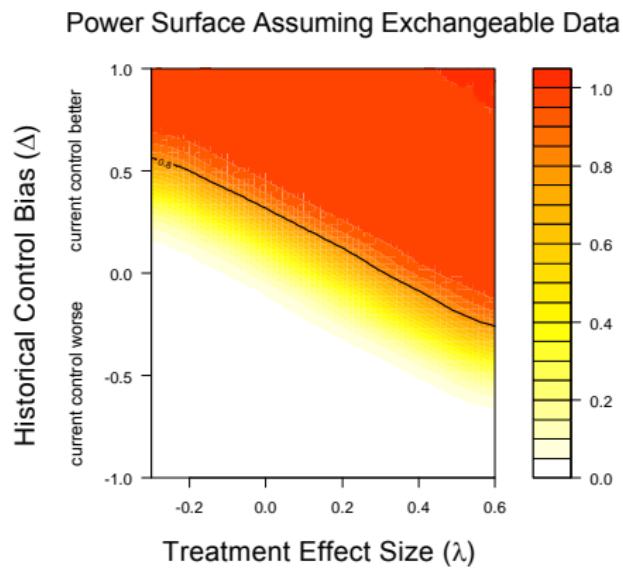
## Exchangeable model: static inference (fixed shrinkage)



## Pocock model: empirical Bayesian inference



# Power Surface: Dynamic Borrowing



## Applying Meta-Analytic-Predictive Priors with the R Bayesian Evidence Synthesis Tools

Sebastian Weber, Yue Li, John W. Seaman III, Tomoyuki Kakizume, Heinz Schmidli

### Abstract

Use of historical data in clinical trial design and analysis has shown various advantages such as reduction of number of subjects and increase of study power. The metaanalytic-predictive (MAP) approach accounts with a hierarchical model for between-trial heterogeneity in order to derive an informative prior from historical data. In this paper, we introduce the package RBesT (R Bayesian evidence synthesis tools) which implements the MAP approach with normal (known sampling standard deviation), binomial and Poisson endpoints. The hierarchical MAP model is evaluated by Markov chain Monte Carlo (MCMC). The MCMC samples representing the MAP prior are approximated with parametric mixture densities which are obtained with the expectation maximization algorithm. The parametric mixture density representation facilitates easy communication of the MAP prior and enables fast and accurate analytical procedures to evaluate properties of trial designs with informative MAP priors. The paper first introduces the framework of robust Bayesian evidence synthesis in this setting and then explains how RBesT facilitates the derivation and evaluation of an informative MAP prior from historical control data. In addition we describe how the meta-analytic framework relates to further applications including probability of success calculations.

Files:

 [Paper](#) [R package \(RBesT\)](#) [R replication code](#)

Published:

Nov 30, 2021

You can help us to improve our editorial management system by reporting possible issues [\[read more\]](#).

### How to Cite

Weber, S., Li, Y., Seaman III, J. W., Kakizume, T., & Schmidli, H. (2021). Applying Meta-Analytic-Predictive Priors with the R Bayesian Evidence Synthesis Tools. *Journal of Statistical Software*, 100(19), 1-32.  
<https://doi.org/10.18637/jss.v100.i19>

[More Citation Formats](#) ▾[Download Citation](#) ▾

### License Information



Article: [Creative Commons Attribution License \(CC-BY\)](#)  
Software: [GPL General Public License version 2 or version 3](#) or a [GPL-compatible license](#).

# Multi-source exchangeability models (MEMs)

Bayesian models that allow for multiple "sources" of exchangeability

- Asymmetric settings (primary & supple cohorts); Kaizer et al. 2017, [\*Biostatistics\*](#)
- Symmetric settings (all cohorts primary) and sequential design; Hobbs and Landin 2018, [\*Stat in Med\*](#)
- Adaptive Platform Design; Kaizer et al. 2018, [\*Biometrics\*](#)
- Frequentist Trial Operating Characteristics; Kaizer et al. 2019, [\*JCO Precision Oncology\*](#)
- Open-source statistical software with the *Basket* package; Kane et al. 2020, [\*The R Journal\*](#)
- Multiple Indication Design Criteria; Kaizer et al. 2021, [\*SMMR\*](#)
- False Discovery Control; Zabor et al. 2022, [\*Clinical Trials\*](#)
- Sequential Master protocol; Kaizer et al. 2022, [\*PLOS One\*](#)

# WHO CAN BE AVERAGED?

## PRECISION MEDICINE FROM THE PERSPECTIVE OF DATA = ASCERTAINING STATISTICAL EXCHANGEABILITY

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<sup>1</sup>  | Rick Landin<sup>2</sup>

<sup>1</sup>Department of Quantitative Health Sciences and the Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio 44195

<sup>2</sup>La Jolla Pharmaceutical Company, San Diego, California 92121

Correspondence  
Brian P. Hobbs, Department of Quantitative Health Sciences and The Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH 44195.  
Email: BHobbs@gmail.com

Precision medicine endeavors to conform therapeutic interventions to the individuals being treated. Implicit to the concept of precision medicine is heterogeneity of treatment benefit among patients and patient subpopulations. Thus, precision medicine challenges conventional paradigms of clinical translational which have relied on estimates of population-averaged effects to guide clinical practice. Basket trials comprise a class of experimental designs used to study solid malignancies that are devised to evaluate the effectiveness of a therapeutic strategy among patients defined by the presence of a particular drug target (often a genetic mutation) rather than a particular tumor histology. Acknowledging the potential for differential effectiveness on the basis of traditional criteria for cancer subtyping, evaluations of treatment effectiveness are con-

The  Journal

### Analyzing Basket Trials under Multisource Exchangeability Assumptions

Michael J. Kane  
Yale University

Nan Chen  
The University  
of Texas

Alexander M. Kaiser  
University of Colorado

Xun Jiang  
Amgen Inc.

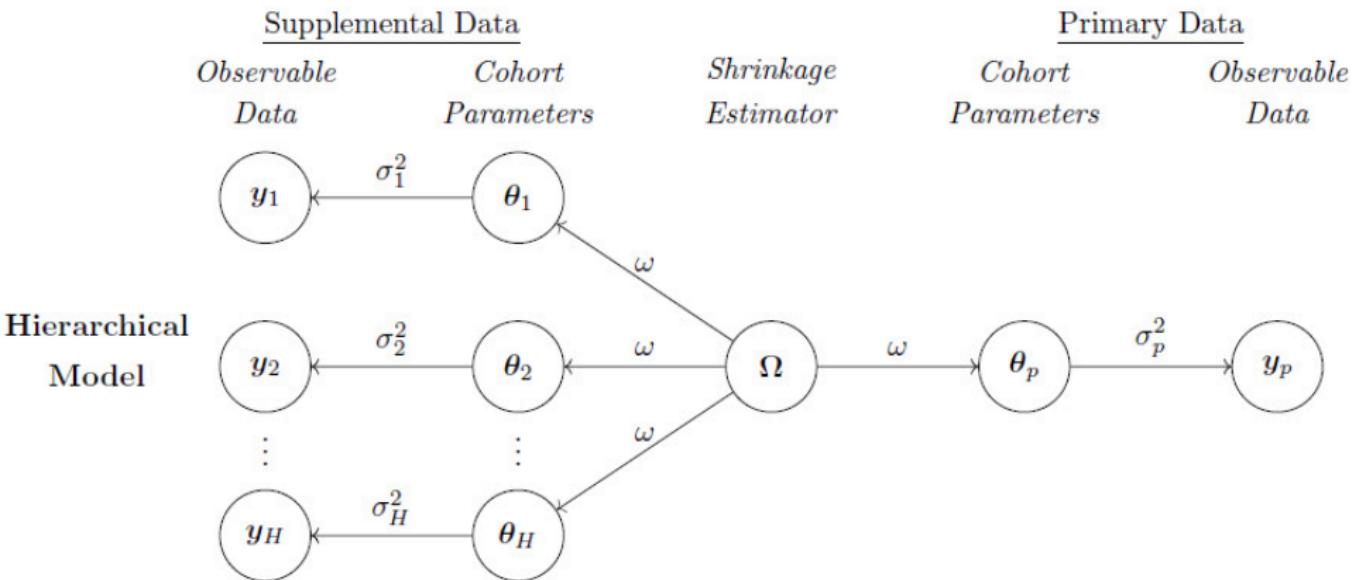
H. Amy Xia  
Amgen Inc.

Brian P. Hobbs  
Cleveland Clinic

#### Abstract

Basket designs are prospective clinical trials that are devised with the hypothesis that the presence of selected molecular features determine a patient's subsequent response to a particular "targeted" treatment strategy. Basket trials are designed to enroll multiple clinical subpopulations to which it is assumed that the therapy in question offers beneficial efficacy in the presence of the targeted molecular profile. The treatment, however, may not offer acceptable efficacy to all subpopulations enrolled. Moreover, for rare dis-

# Conventional Hierarchical Models are limited!



# Bayesian hierarchical modeling based on multisource exchangeability

ALEXANDER M. KAIZER, JOSEPH S. KOOPMEINERS\*

*Division of Biostatistics, University of Minnesota, A460 Mayo Building, MMC 303 420 Delaware St. SE,  
Minneapolis, MN 55455, USA*

koopm007@umn.edu

BRIAN P. HOBBS

*The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd. Houston, TX 77030, USA*

## SUMMARY

Bayesian hierarchical models produce shrinkage estimators that can be used as the basis for integrating supplementary data into the analysis of a primary data source. Established approaches should be considered limited, however, because posterior estimation either requires prespecification of a shrinkage weight for each source or relies on the data to inform a single parameter, which determines the extent of influence or shrinkage from all sources, risking considerable bias or minimal borrowing. We introduce multisource exchangeability models (MEMs), a general Bayesian approach for integrating multiple, potentially non-exchangeable, supplemental data sources into the analysis of a primary data source. Our proposed modeling framework yields source-specific smoothing parameters that can be estimated in the presence of the data to facilitate a dynamic multi-resolution smoothed estimator that is asymptotically consistent while reducing the dimensionality of the prior space. When compared with competing Bayesian hierarchical modeling strategies, we demonstrate that MEMs achieve approximately 2.2 times larger median effective supplemental sample size when the supplemental data sources are exchangeable as well as a 56% reduction in bias when there is heterogeneity among the supplemental sources. We illustrate the application of MEMs using a recently completed randomized trial of very low nicotine content cigarettes, which resulted in a 30% improvement in efficiency compared with the standard analysis.

**Keywords:** Bayesian hierarchical modeling; Heterogeneous sources of data; Multisource smoothing; Supplementary data.

# Kaizer, Koopmeiners, Hobbs (2017) Bayesian hierarchical modeling based on multi-source exchangeability. Biostatistics

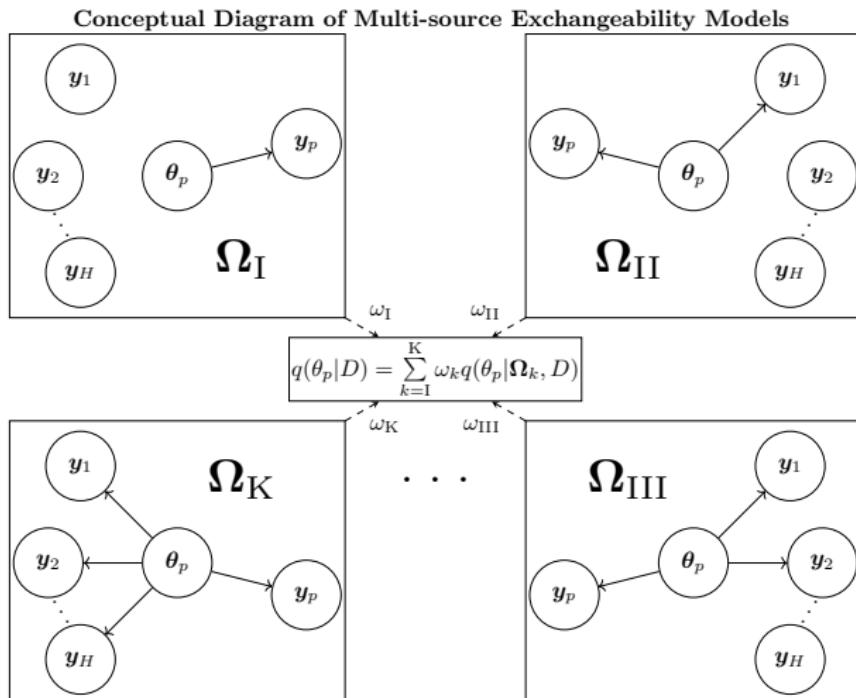


Fig. 1: Each MEM is a combination of supplemental sources assumed exchangeable with the primary cohort in order to estimate the parameters of interest,  $\theta_p$ , and is contained within each box for  $\Omega_k$ . Within a box the solid arrows  $\theta_p$  and the observables,  $y_h$ , represent which supplemental sources are assumed exchangeable with the primary cohort within the given MEM.

# Case Studies

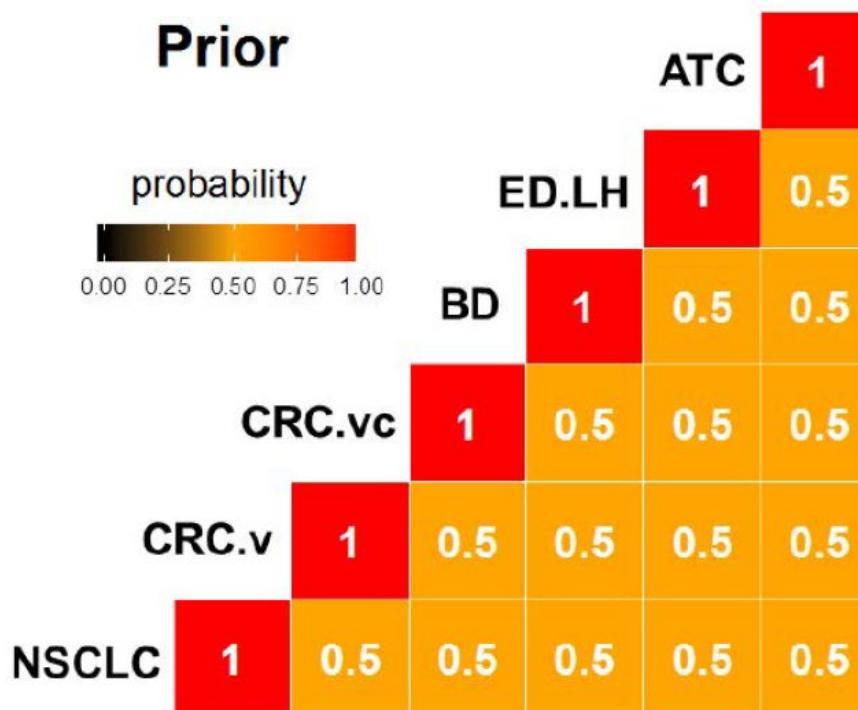
## Case Study: Vemurafenib non-melanoma basket trial

Baskets	Enrolled	Evaluable	Responders	Posterior probability $Pr(\pi > 0.15)$ based on response only
NSCLC	20	19	8	<b>0.998</b>
CRC (vemu)	10	10	0	<b>0.068</b>
CRC (vemu + cetu)	27	26	1	<b>0.039</b>
Bile Duct	8	8	1	<b>0.472</b>
ECD or LCH	18	14	6	<b>0.995</b>
ATC	7	7	2	<b>0.847</b>

- Bayesian Posterior Probability  $Pr(\pi > 0.15 | Data) > \theta$ , with  $\theta$  fixed to control type I error at 0.10
- data reported in article: "Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations," *NEJM* (2015)

# Bayesian Basket Trial Design with Exchangeability Monitoring

Brian P. Hobbs<sup>1</sup> and Rick Landin<sup>2</sup>



# Case Study Analysis: Vemurafenib non-melanoma basket trial

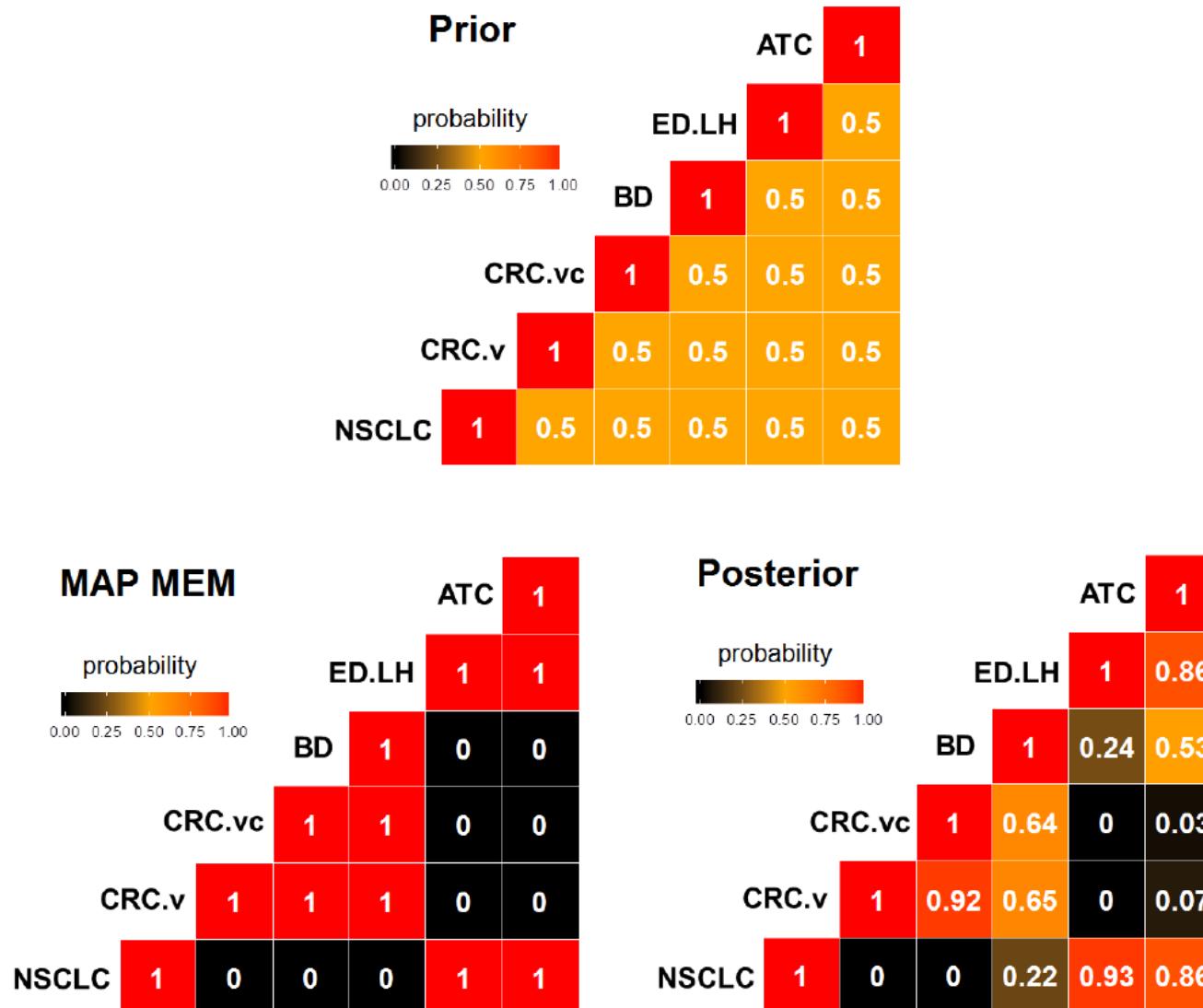


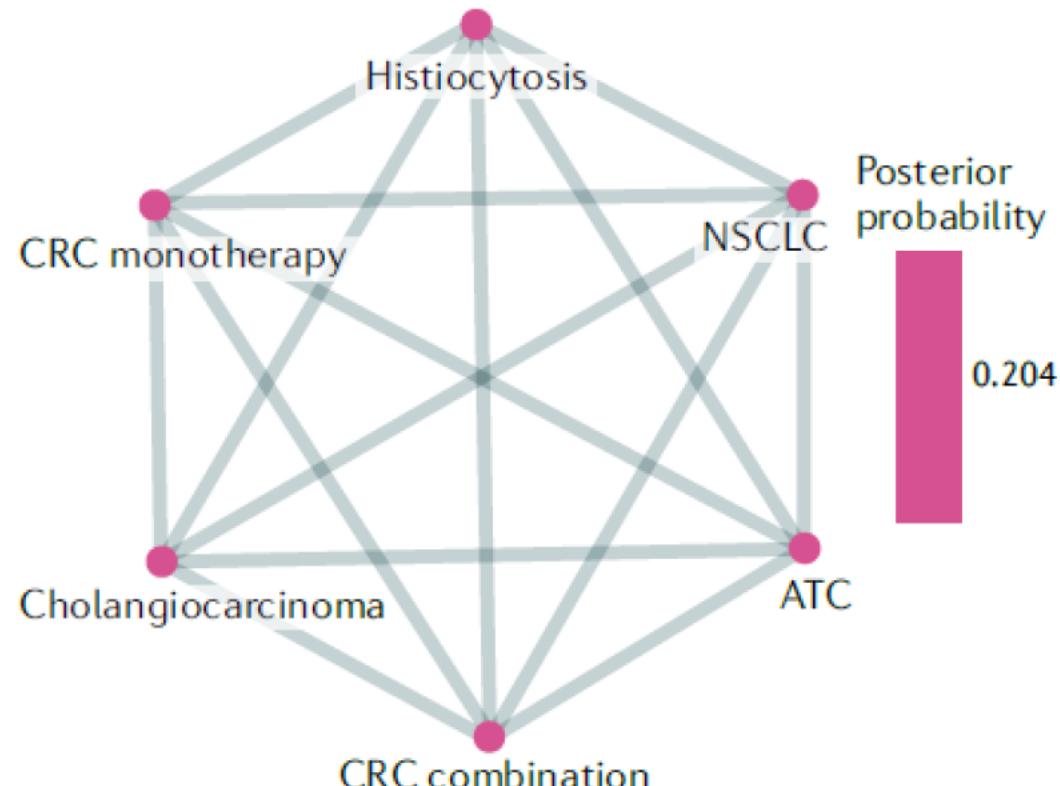
Figure 2. Prior, MAP, and PEP that result from Bayesian inference using the observed vemurafenib basket trial data

# Histology-agnostic drug development – considering issues beyond the tissue

Roberto Carmagnani Pestana, Shiraj Sen, Brian P. Hobbs & David S. Hong 

*Nature Reviews Clinical Oncology* **17**, 555–568(2020)

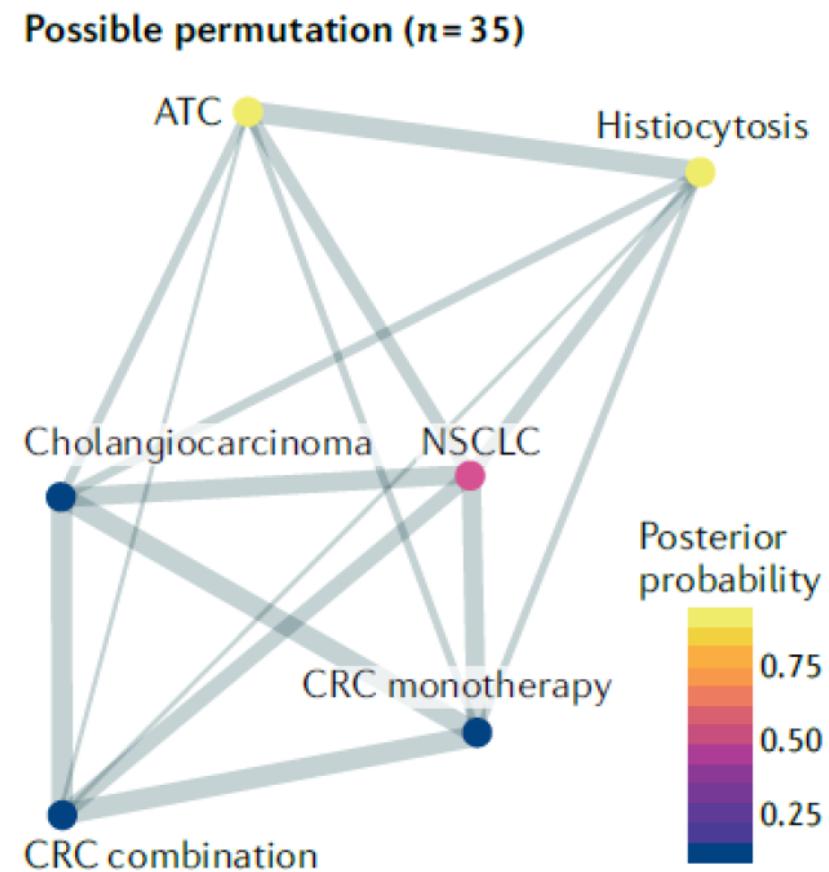
**Baseline assumptions ( $n = 0$ )**



# Histology-agnostic drug development – considering issues beyond the tissue

Roberto Carmagnani Pestana, Shiraj Sen, Brian P. Hobbs & David S. Hong 

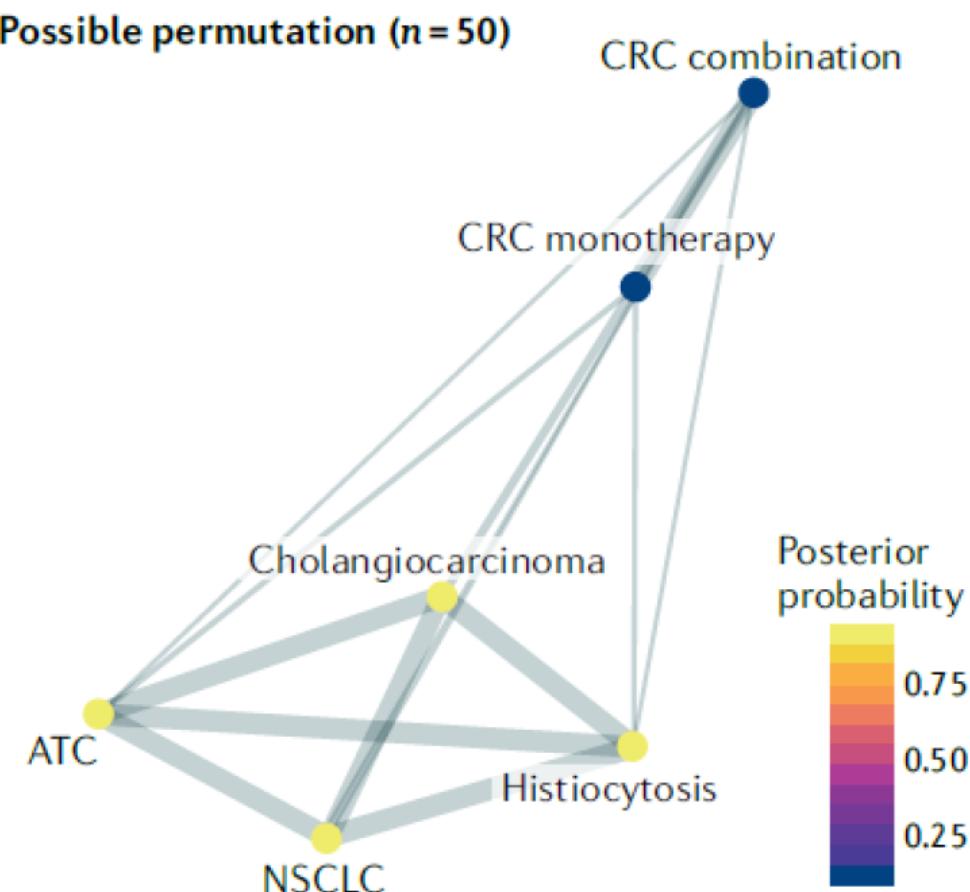
*Nature Reviews Clinical Oncology* **17**, 555–568(2020)



## Histology-agnostic drug development – considering issues beyond the tissue

Roberto Carmagnani Pestana, Shiraj Sen, Brian P. Hobbs & David S. Hong 

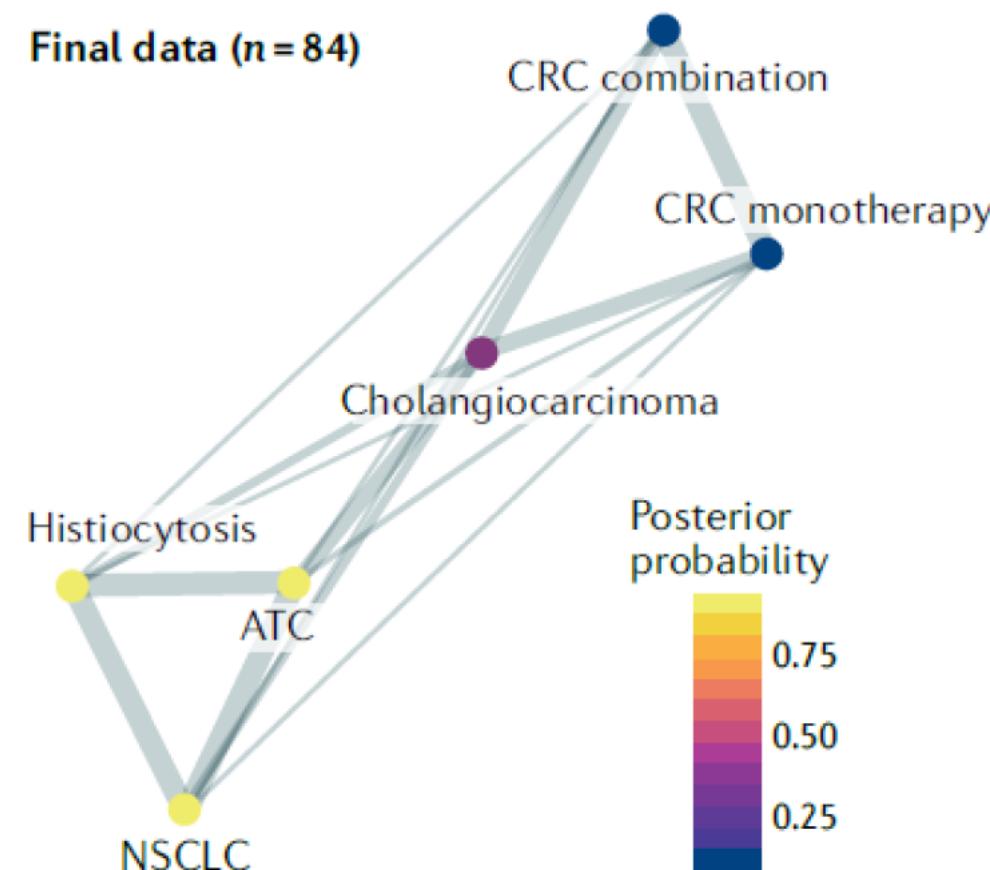
*Nature Reviews Clinical Oncology* **17**, 555–568(2020)



# Histology-agnostic drug development – considering issues beyond the tissue

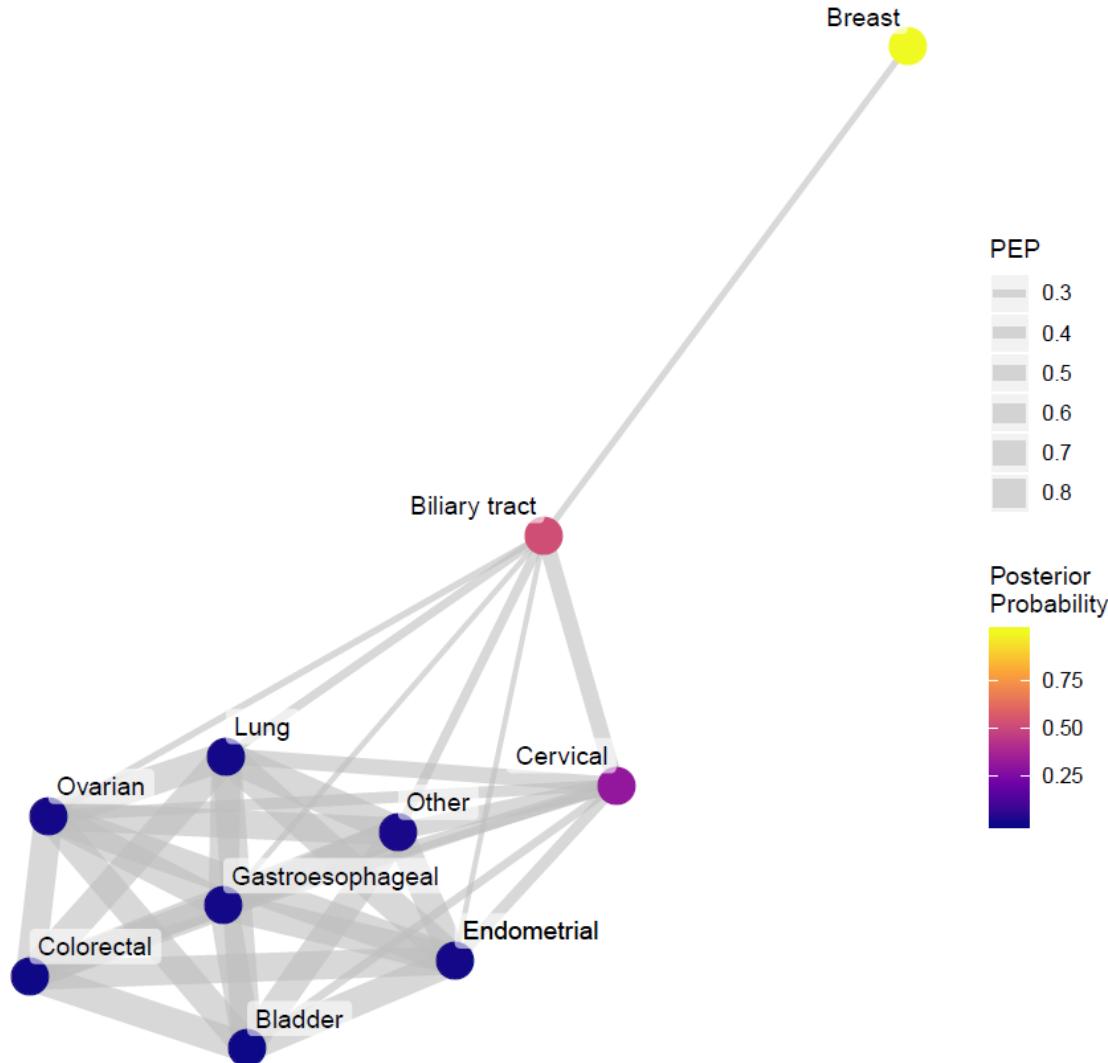
Roberto Carmagnani Pestana, Shiraj Sen, Brian P. Hobbs & David S. Hong 

*Nature Reviews Clinical Oncology* **17**, 555–568(2020)



# Network graph of SUMMIT results

```
plot_pep_graph(nerat_basket, pep_cutoff = 0.25)
```



# Bayesian hierarchical modeling based on multisource exchangeability

Alexander M Kaizer, Joseph S Koopmeiners ✉, Brian P Hobbs

*Biostatistics*, Volume 19, Issue 2, 1 April 2018, Pages 169–184, <https://doi.org/10.1093/biostatistics/kxx051>.

Let  $L$  represent the likelihood, given the data for  $\Omega_k$ ,  $\Theta = (\theta_p, \theta_1, \dots, \theta_H)$ , and  $\pi(\Theta|\Omega_k)$  denote the prior density of  $\Theta$  under  $\Omega_k$ . In the context of MEMs, the integrated marginal likelihood for a particular MEM, given the data are obtained by averaging the likelihood over the posterior distribution for the vector of all model parameters of interest,

$$p(D|\Omega_k) = \int L(\Theta|D, \Omega_k) \pi(\Theta|\Omega_k) d\Theta. \quad (2.1)$$

The posterior model weights for each MEM are given by

$$\omega_k = p(\Omega_k|D) = \frac{p(D|\Omega_k)\pi(\Omega_k)}{\sum_{j=1}^K p(D|\Omega_j)\pi(\Omega_j)}, \quad (2.2)$$

where  $\pi(\Omega_k)$  is the prior probability that  $\Omega_k$  is the true model. The marginal posterior distribution, given the observable data  $D$  to be used for inference on  $\theta_p$  is the weighted average using the posterior model

# Bayesian hierarchical modeling based on multisource exchangeability

Alexander M Kaizer, Joseph S Koopmeiners ✉, Brian P Hobbs

*Biostatistics*, Volume 19, Issue 2, 1 April 2018, Pages 169–184, <https://doi.org/10.1093/biostatistics/kxx051>.

*Bayesian hierarchical modeling based on MEMs*

173

weights of the  $K$  MEM posteriors,  $p(\theta_p | \Omega_k, D)$ :

$$p(\theta_p | D) = \sum_{k=1}^K \omega_k p(\theta_p | \Omega_k, D). \quad (2.3)$$

Unfortunately, BMA quickly becomes highly parameterized, as the number of models grows exponentially with the number of supplementary sources ( $K = 2^H$ ), and prior specification over a model space of large size is problematic. Fernández and others (2001) noted that posterior model weights can be very sensitive to the specification of priors in the model, especially in the absence of strong prior knowledge. In the analysis of limited data obtained from a clinical study, these issues with the conventional BMA approach become critical and motivate our proposed approach. With MEMs, the supplemental sources are assumed to be distinct and independent, therefore we can specify priors with respect to sources instead of models,  $\pi(\Omega_k) = \pi(S_1 = s_{1,k}, \dots, S_H = s_{H,k}) = \pi(S_1 = s_{1,k}) \times \dots \times \pi(S_H = s_{H,k})$ . This results in drastic dimension reduction in that it necessitates the specification of only  $H$  source-specific prior inclusion probabilities in place of  $2^H$  prior model probabilities comprising the entire model space. In Section 3.2, we propose prior weights for the source-inclusion probabilities that result in consistent posterior model weights and yield desirable small sample properties, as evaluated by simulation in Section 4. In contrast, similarly constructed prior weights on the models did not result in consistent posterior model weights.

# Bayesian Basket Trial Design with Exchangeability Monitoring

Brian P. Hobbs<sup>1</sup> and Rick Landin<sup>2</sup>

hyperparameters  $a$  and  $b$ , upon having observed successes  $S = \{S_1, \dots, S_J\}$ , Bayes' Theorem yields the following conjugate conditional posterior distribution for the response probability of basket  $j$  representing the Bayesian update of  $p(\pi | S_{(-j)})$  with likelihood  $\text{Bin}(S_j | \pi_j, n_j)$ ,

$$q(\pi_j | S, \Omega_j) \propto \text{Beta}\left(a + \sum_{h=1}^J \Omega_{j,h} S_h, b + \sum_{k=1}^J \Omega_{j,k} (n_k - S_k)\right), \quad (1)$$

where  $\Omega_j$  represents the  $j^{th}$  row of multisource exchangeability matrix  $\Omega$ . Marginal posterior inference with respect to  $\pi_j | S$  averages (1) with respect to the marginal posterior probability of  $G = 2^{J-1}$  possible exchangeability configurations of  $\Omega_j$ . Let  $\omega = \{\omega_1, \dots, \omega_G\}$  denote the collection of vectors each of length  $J$  and with  $j^{th}$  element = 1 that collectively span the sample space of  $\Omega_j$ . The marginal posterior distribution can be represented by a finite mixture density

$$q(\pi_j | S) \propto \sum_{g=1}^G q(\pi_j | S, \Omega_j = \omega_g) Pr(\Omega_j = \omega_g | S), \quad (2)$$

where the posterior probability of exchangeability configuration  $\omega_g$  given the observed data follows from Bayes' Theorem in proportion to the marginal density of the data given  $\omega_g$  and its unconditional prior probability

$$Pr(\Omega_j = \omega_g | S) \propto \frac{m(S_j | \Omega_j = \omega_g, S_{(-j)}) Pr(\Omega_j = \omega_g)}{\sum_{u=1}^G m(S_j | \Omega_j = \omega_u, S_{(-j)}) Pr(\Omega_j = \omega_u)}. \quad (3)$$

# Bayesian Basket Trial Design with Exchangeability Monitoring

Brian P. Hobbs<sup>1</sup> and Rick Landin<sup>2</sup>

Let  $B()$  denote the beta function. Given an exchangeability configuration  $\Omega_j$ , the marginal density of  $S_j$  may be obtained by integrating the likelihood of  $\pi_j | S_j$  with respect to  $p(\pi | S_{(-j)})$ ,

$$m(S_j | \Omega_j, S_{(-j)}) \propto \frac{B\left(a + \sum_{h=1}^J \Omega_{j,h} S_h, b + \sum_{k=1}^J \Omega_{j,k} (n_k - S_k)\right)}{B(a, b)} \times \\ \prod_{i=1}^J \left( \frac{B(a + S_i, b + n_i - S_i)}{B(a, b)} \right)^{1-\Omega_{j,i}}. \quad (4)$$

The Bayesian model is complete given specification of a vector comprising the unconditional prior probabilities of all possible pairwise exchangeability configurations,  $Pr(\Omega)$ , which is challenging given the high dimensionality of the MEM sample domain. Defining  $Pr(\Omega)$  as the product of prior exchangeability probabilities for each unique basket pair, however, reduces the dimension from  $\prod_{k=1}^{J-1} 2^k$  to  $J(J-1)/2$  yielding feasibility and thereby offering an advantage with respect to conventional Bayesian model averaging

$$Pr(\Omega) = Pr(\Omega_{1,2} = 1) \times Pr(\Omega_{1,3} = 1) \times \cdots \times Pr(\Omega_{J-1,J} = 1). \quad (5)$$

By the Kolmogorov definition of conditional probability, the prior exchangeability probabilities for all  $2^{J-1}$  configurations of  $\Omega_j$  in (3) follow from (5) as  $Pr(\Omega_j = \omega) = \prod_{i=1}^J Pr(\Omega_{j,i} = 1)^{I(\omega_i=1)} \times \{1 - Pr(\Omega_{j,i} = 1)\}^{(1-I(\omega_i=1))}$ , where  $I()$  is the indicator function and  $\omega$  represents one vector of length  $J$  within the sample domain of  $\Omega_j$ .

### 2.3. Posterior Probability and Effective Sample Size

The MEM Bayesian model specification facilitates posterior inference with respect to all possible pairwise exchangeability relationships among  $J$  subtypes. The framework facilitates estimation of disjointed subpopulations comprised of meta-subtypes or singleton subtypes and thereby offers additional flexibility when compared to SEM specifications. The uncontrolled basket study considered herein is devised with the intention of testing the hypothesis that the response probability for a targeted intervention exceeds a null value, which we denote  $\pi_0$ , while acknowledging the potential for heterogeneity in effectiveness in accordance with the pre-specified basket partitions. Within the MEM framework, this testing procedure follows from the cumulative density function (cdf) of the marginal posterior distribution (2). Specifically, the posterior probability that  $\pi_j$  exceeds  $\pi_0$  may be computed as the weighted average of cdfs for all possible exchangeability configurations,

$$Pr(\pi_j > \pi_0 | S) = \sum_{g=1}^G Pr(\Omega_j = \omega_g | S) \left\{ 1 - \frac{\int_0^{\pi_0} u^{a+\sum_{h=1}^J \omega_{g,h} S_h - 1} (1-u)^{b+\sum_{k=1}^J \omega_{g,k} (n_k - S_k) - 1} du}{B(a + \sum_{h=1}^J \omega_{g,h} S_h, b + \sum_{k=1}^J \omega_{g,k} (n_k - S_k) - 1)} \right\}. \quad (6)$$

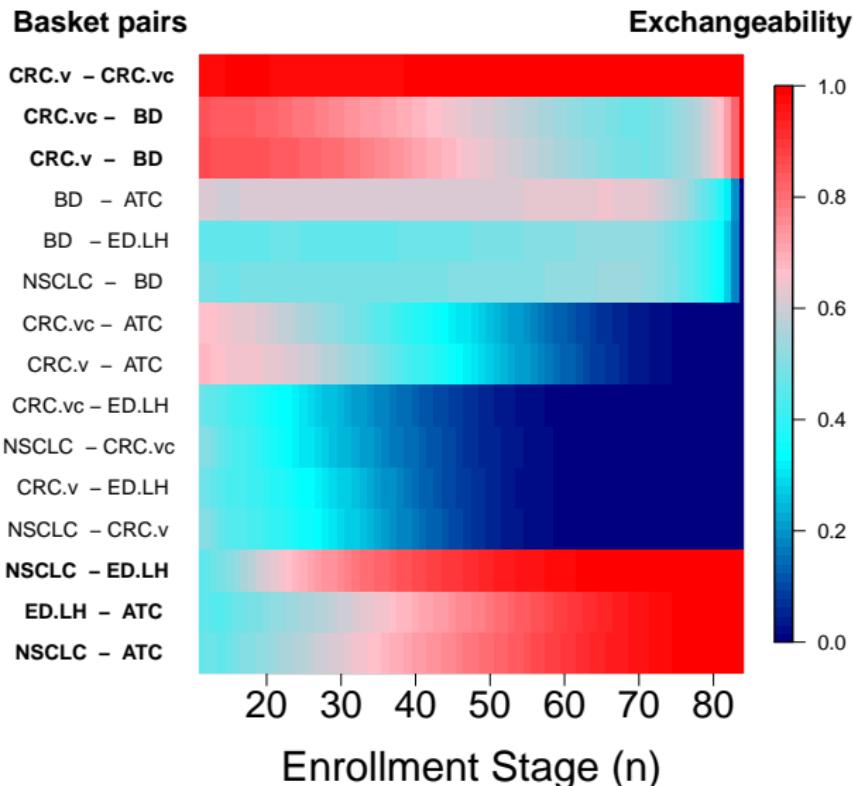
Measurement of the extent to which information has been shared across subtypes in the context of a Bayesian analysis is best characterized by the effective sample size (ESS) of the resultant posterior distribution [see e.g. 17, 18, 19, 10]. The ESS quantifies information content in relation to the number of observables that would be required to obtain the level of posterior precision achieved by the candidate posterior distribution when analyzed using a vague “reference” or maximum entropy prior. Given a specific MEM, the ESS of the conditional posterior in (1) can be derived as

$$ESS(\Omega_j) = a + b + \sum_{h=1}^J \Omega_{j,h} n_h. \quad (7)$$

The marginal posterior ESS may be approximated by the weighted average of the individual MEM-specific values of ESS as  $ESS = \sum_{g=1}^G Pr(\Omega_j = \omega_g | S) ESS(\Omega_j)$ . Because the initial beta prior distribution carries the effective information of  $a + b$  patients, the MEM model facilitates a minimum ESS of  $a + b + n_j$  for posterior inference of  $\pi_j | S$ , which is

# **Sequential Design based on Exchangeability Monitoring with MEM**

# Permutation Study: Vemurafenib non-melanoma basket trial



# Permutation Study: Vemurafenib non-melanoma basket trial

