Historical controls in clinical trials: the meta-analytic predictive approach applied to over-dispersed count data

Sandro Gsteiger, Beat Neuenschwander, and Heinz Schmidli

Novartis Pharma AG

Bayes Pharma, Aachen, May 2012

Table of contents

Table of contents

Introduction

- Relevant information always available
- Informal use in trial design and planning is standard
  - Endpoints
  - Design options
  - Assumptions for sample size calculation
- Increasing interest in formal use also for analysis
  - Methodological: bias model, power priors, meta-analytic predictive priors (hierarchical modeling), commensurate priors. Approaches similar in spirit; closely related formalisms.
  - Practical (orphan indications, medical devices, pediatrics)
- Key issue: proper discounting of historical data
  - How much do we know?
  - "Worth" how many patients, n^*?

Introduction

- Relevant information always available
- Informal use in trial design and planning is standard
  - Endpoints
  - Design options
  - Assumptions for sample size calculation
- Increasing interest in formal use also for analysis
  - Methodological: bias model, power priors, meta-analytic predictive priors (hierarchical modeling), commensurate priors. Approaches similar in spirit; closely related formalisms.
  - Practical (orphan indications, medical devices, pediatrics)
- Key issue: proper discounting of historical data
  - How much do we know?
  - "Worth" how many patients, n^*?

Introduction

- Relevant information always available
- Informal use in trial design and planning is standard
  - Endpoints
  - Design options
  - Assumptions for sample size calculation
- Increasing interest in formal use also for analysis
  - Methodological: bias model, power priors, meta-analytic predictive priors (hierarchical modeling), commensurate priors. Approaches similar in spirit; closely related formalisms.
  - Practical (orphan indications, medical devices, pediatrics)
- Key issue: proper discounting of historical data
  - How much do we know?
  - ”Worth” how many patients, n^*?

Objective of using historical data

- More powerful analysis by using more information
  - Either augment the data (n *vs* n + n^*)
  - Or replace control patients (n *vs* (n - n^*) + n^*)
  - Example: study active *vs* control, groups size $n = 20$, historical data worth $n^* = 10$ controls.

    Standard: new trial "20 *vs* 20"; analysis: data worth 40 patients.
    Augment: new trial "20 *vs* 20"; analysis: data worth 50 patients.
    Replace: new trial "20 *vs* 10"; analysis: data worth 40 patients.

- So how big should n^* be?
- Extreme positions on use of historical control data
  - Complete pooling: future and past response are equal
  - No borrowing: separate analysis
- Both positions are unrealistic, need structured approach for compromize (proper discounting)

Objective of using historical data

- More powerful analysis by using more information
  - Either augment the data (n *vs* n + n^*)
  - Or replace control patients (n *vs* (n - n^*) + n^*)
  - Example: study active *vs* control, groups size $n = 20$, historical data worth $n^* = 10$ controls.

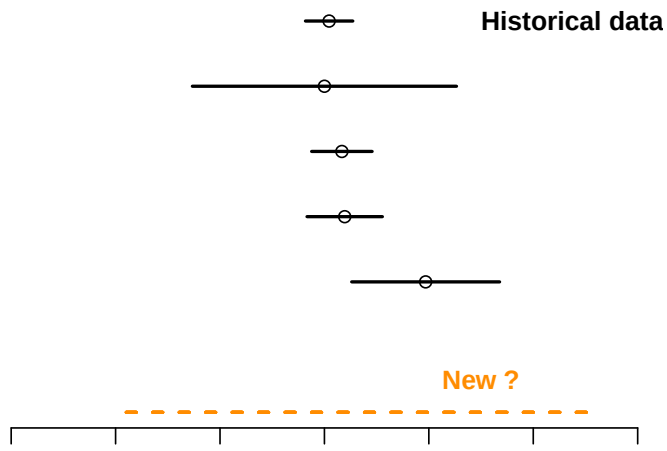
    Standard: new trial "20 *vs* 20"; analysis: data worth 40 patients.
    Augment: new trial "20 *vs* 20"; analysis: data worth 50 patients.
    Replace: new trial "20 *vs* 10"; analysis: data worth 40 patients.

- So how big should n^* be?
- Extreme positions on use of historical control data
  - Complete pooling: future and past response are equal
  - No borrowing: separate analysis
- Both positions are unrealistic, need structured approach for compromize (proper discounting)

Design issues

Trying to predicting the future from knowing the past

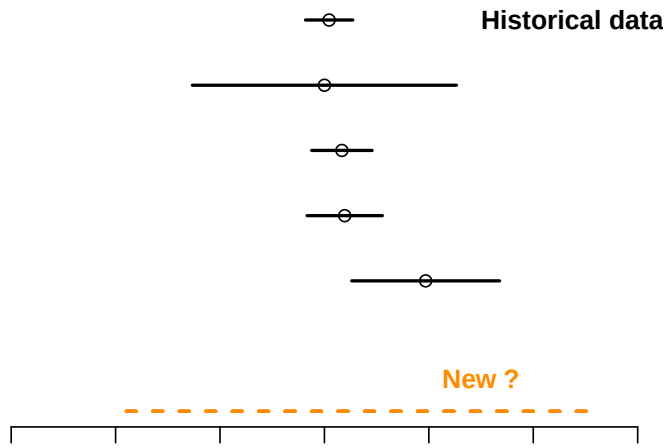
- Difficult problem!
- In indication ABC, placebo response is about ...
- How well do we know the past, how much can we predict?
- Sometimes prediction with confidence...
- ... or with large uncertainty.



Design issues

Trying to predicting the future from knowing the past

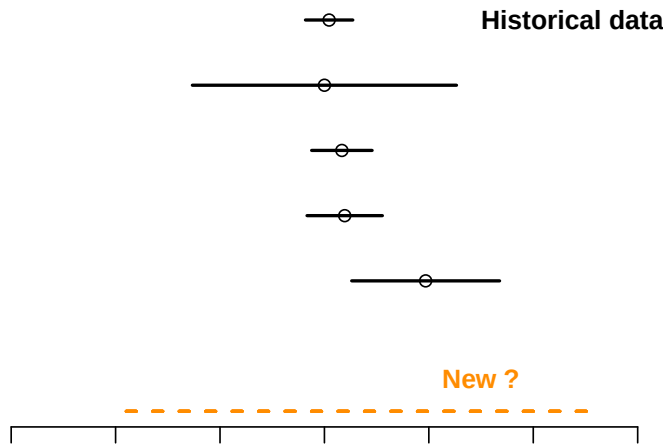
- Difficult problem!
- In indication ABC, placebo response is about ...
- How well do we know the past, how much can we predict?
- Sometimes prediction with confidence...
- ... or with large uncertainty.

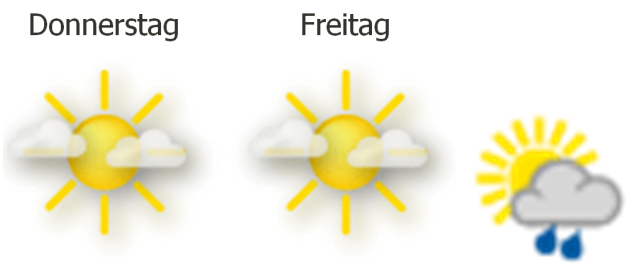


Design issues

Trying to predicting the future from knowing the past

- Difficult problem!
- In indication ABC, placebo response is about ...
- How well do we know the past, how much can we predict?
- Sometimes prediction with confidence...
- ... or with large uncertainty.





Analysis issues

- Analysis: fundamentally, two philosophies
  1. Joint analysis of historical and new data at end of new study (with parameter of interest effect in new trial)
  2. Derive informative prior upfront
- Mathematically both approaches equivalent (in principle)
- Will focus on construction of historical data prior
  - Conceptually clear separation of information sources
  - Quantification of information upfront (n^*)
- Challenges for communication with clinical teams
  - Two sources of information
  - Possibility of conflict

Analysis issues

- Analysis: fundamentally, two philosophies
  1. Joint analysis of historical and new data at end of new study (with parameter of interest effect in new trial)
  2. Derive informative prior upfront
- Mathematically both approaches equivalent (in principle)
- Will focus on construction of historical data prior
  - Conceptually clear separation of information sources
  - Quantification of information upfront (n^*)
- Challenges for communication with clinical teams
  - Two sources of information
  - Possibility of conflict

Analysis issues

- Analysis: fundamentally, two philosophies
  1. Joint analysis of historical and new data at end of new study (with parameter of interest effect in new trial)
  2. Derive informative prior upfront
- Mathematically both approaches equivalent (in principle)
- Will focus on construction of historical data prior
  - Conceptually clear separation of information sources
  - Quantification of information upfront (n^*)
- Challenges for communication with clinical teams
  - Two sources of information
  - Possibility of conflict

Extension to MAP approach studied here

- MAP was introduced for (approx.) normal endpoints
- Objective here: apply to over-dispersed counts
  - Consider two data sources: summary and individual level
  - Summary level: literature, trial repositories etc.
  - Individual level: in-house trials
- Examples of over-dispersed counts in pharma
  - Multiple sclerosis: lesion counts on MRI scans, number of relapses
  - COPD: number of exacerbations
  - Number of AEs within a patient
  - ...

Extension to MAP approach studied here

- MAP was introduced for (approx.) normal endpoints
- Objective here: apply to over-dispersed counts
  - Consider two data sources: summary and individual level
  - Summary level: literature, trial repositories etc.
  - Individual level: in-house trials
- Examples of over-dispersed counts in pharma
  - Multiple sclerosis: lesion counts on MRI scans, number of relapses
  - COPD: number of exacerbations
  - Number of AEs within a patient
  - ...

Table of contents

The meta-analytic predictive approach

Historical trials, $h = 1, \ldots, H$:

$$y_h \sim F(\theta_h; n_h, \ldots)$$

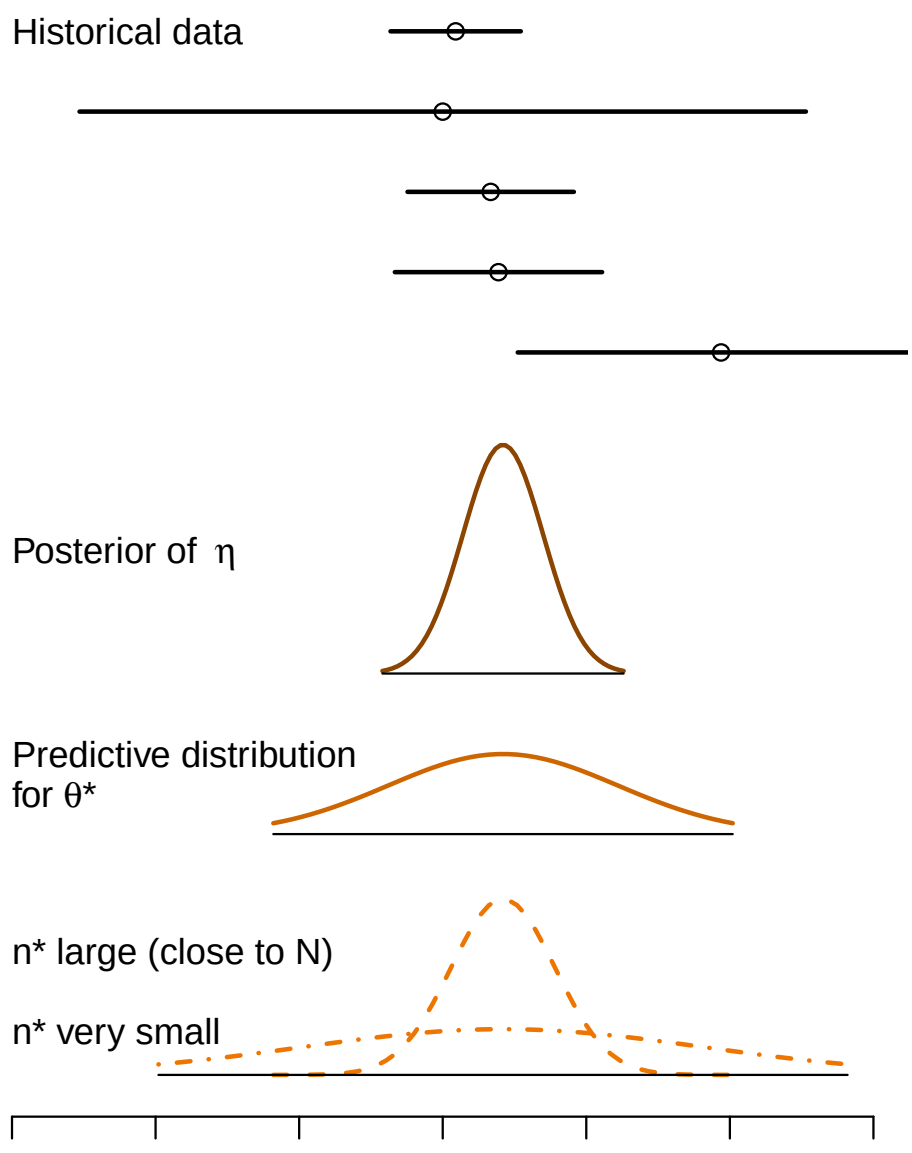
New trial: $Y^* \sim F(\theta^*; \ldots)$

Exchangeability assumption:

$$\theta_1, \ldots, \theta_H, \theta^* | \eta, \tau \overset{\text{iid}}{\sim} N(\eta, \tau^2)$$

Prior for new trial mean:
$\pi(\theta^* | y_1, \ldots, y_H)$

"Worth" n^* patients
(prior effective sample size).



Exchangeability - the key assumption

- Formally: $\pi(\theta_1, \ldots, \theta_H, \theta^*)$ is permutation invariant
  - Difficult to communicate to clinical teams
  - "For any two randomly selected trials, no specific ordering between underlying means expected."
- May apply after covariate adjustment (meta-regression).
- In practice, not always considered carefully enough
  - The new trial mean θ^* is part of it
  - Changes in trial design may lead to new round of MAP (e.g. inclusion criteria)
  - Trial design: iterative process; MA: time-consuming
- Potential source of prior-data conflict if violated

Exchangeability - the key assumption

- Formally: $\pi(\theta_1, \ldots, \theta_H, \theta^*)$ is permutation invariant
  - Difficult to communicate to clinical teams
  - "For any two randomly selected trials, no specific ordering between underlying means expected."
- May apply after covariate adjustment (meta-regression).
- In practice, not always considered carefully enough
  - The new trial mean θ^* is part of it
  - Changes in trial design may lead to new round of MAP (e.g. inclusion criteria)
  - Trial design: iterative process; MA: time-consuming
- Potential source of prior-data conflict if violated

MAP with over-dispersed counts endpoint

Sampling model

- Negative binomial distribution, $NB(\mu, \kappa)$
- Trial specific means μ_h and common dispersion κ

For individual level data

- Y_{ih} the count for patient i in study h
- Implement NB as Gamma-Poisson mixture

$$\begin{aligned} Y_{ih}|\lambda_{ih} &\sim Poi(\lambda_{ih}), \\ \lambda_{ih}|\mu_h, \kappa &\sim Gamma(1/\kappa, 1/(\kappa\mu_h)). \end{aligned}$$

- Then

$$\mathrm{E}[Y_{ih}] = \mu_h, \qquad \mathrm{Var}(Y_{ih}) = \mu_h + \kappa\mu_h^2.$$

MAP with over-dispersed counts (ctd)

Summary level data

- Only observed means $\hat{\mu}_h = \sum_i Y_{ih}/n_h$ given
- This sum of NBs is again NB,

$$\sum_i Y_{ih} \sim NB(n_h\mu_h, \kappa/n_h)$$

Exchangeable log-means, $\theta_h = \log(\mu_h)$,

$$\theta_1, \ldots, \theta_H, \theta^* | \eta, \tau \sim N(\eta, \tau^2).$$

Prior specification

Complete the model with prior for (κ, η, τ). For example

- Uninformative $\pi(\eta)$ often appropriate (well determined)
- Exponential distribution for κ, $\pi(\kappa) = e^{-\kappa/m}/m$
  - Maximum entropy distribution under constraint $\mathrm{E}(\kappa) = m$
- τ needs careful consideration!
  - If only few trials, τ not well determined by data.
  - Some often used "non-informative" priors such as $IG(\varepsilon, \varepsilon)$ have problematic properties (Gelman (2006)).
  - Must decide on plausible prior; needs good judgement.
  - Suggest half-normals with most mass on meaningful range, i.e. small probability for τ leading to no borrowing.

Prior specification

Complete the model with prior for (κ, η, τ). For example

- Uninformative $\pi(\eta)$ often appropriate (well determined)
- Exponential distribution for κ, $\pi(\kappa) = e^{-\kappa/m}/m$
  - Maximum entropy distribution under constraint $\mathrm{E}(\kappa) = m$
- τ needs careful consideration!
  - If only few trials, τ not well determined by data.
  - Some often used "non-informative" priors such as $IG(\varepsilon, \varepsilon)$ have problematic properties (Gelman (2006)).
  - Must decide on plausible prior; needs good judgement.
  - Suggest half-normals with most mass on meaningful range, i.e. small probability for τ leading to no borrowing.

Prior effective sample size

Analysis of new trial using the historical data: take posterior predictive distribution $\pi(\theta^*|y)$ as prior for placebo group in new study.

If $\pi(\theta^*|y)$ approximate normal, prior effective sample size (Neuenschwander et al. (2010))

$$n^* = N\frac{\text{Var}(\theta^*|y, \tau = 0)}{\text{Var}(\theta^*|y)}.$$

- Ballpark figure; tool for design of new study.
- O.k. in the NB example later on., but other cases may need different approximations; e.g. if Beta(a, b), then $n^* = a + b$.

Prior effective sample size

Analysis of new trial using the historical data: take posterior predictive distribution $\pi(\theta^*|y)$ as prior for placebo group in new study.

If $\pi(\theta^*|y)$ approximate normal, prior effective sample size (Neuenschwander et al. (2010))

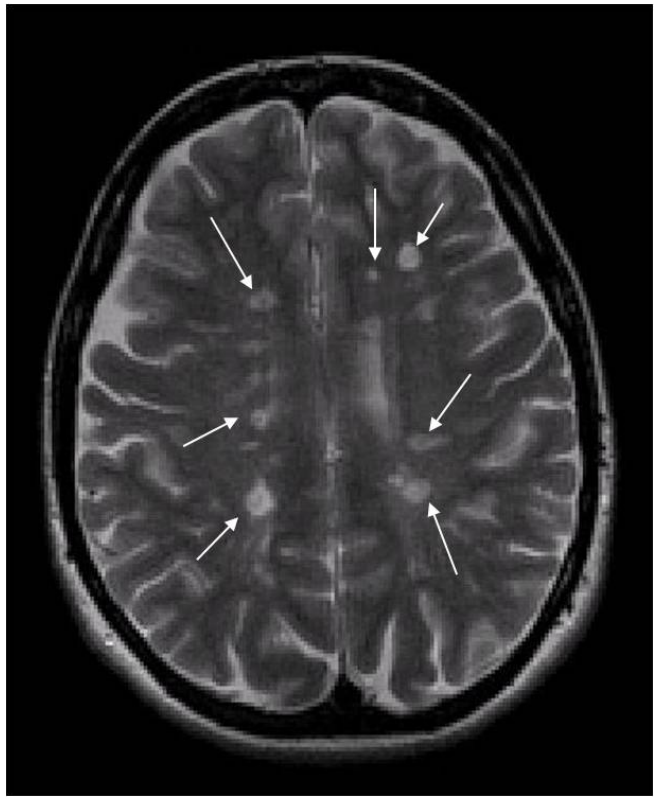
$$n^* = N\frac{\text{Var}(\theta^*|y, \tau = 0)}{\text{Var}(\theta^*|y)}.$$

- Ballpark figure; tool for design of new study.
- O.k. in the NB example later on., but other cases may need different approximations; e.g. if Beta(a, b), then $n^* = a + b$.

Table of contents

Example: lesion counts in multiple sclerosis

- Progressive, degenerative disease of the central nervous system
- Inflammation and tissue damage in the brain
- Lesions on MRI scans used in diagnosis and monitoring of disease evolution
- Lesions serve as (early) markers in clinical trials
- Lesion counts typically over-dispersed



Data overview (placebo patients)

Baseline enrichment

- Enrol only patients with ≥ 1 lesion at baseline
- Minimal disease activity (more homogeneous population)

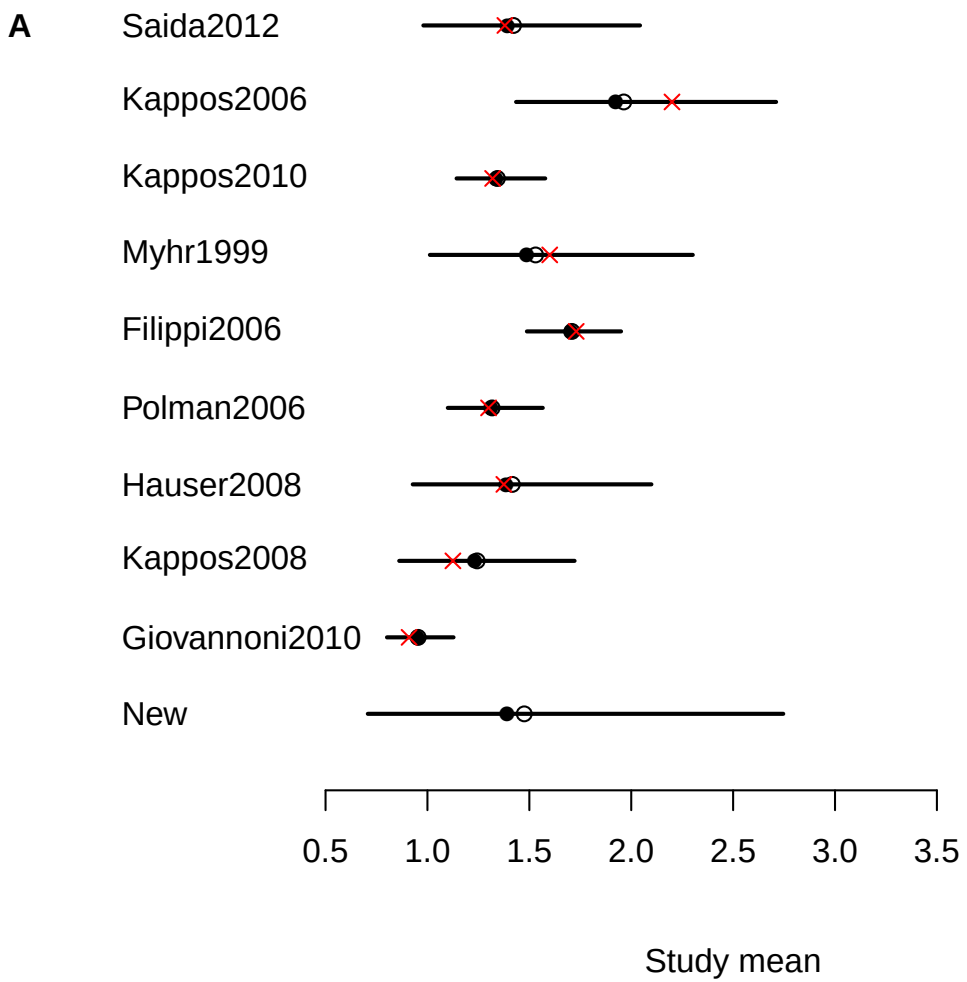
		A		**B**	
	Study	n_h	$\hat{\mu}_h$	n_h	$\hat{\mu}_h$
In-house	Kappos2006	81	2.2	41	3.29
	Kappos2010	373	1.32	129	2.91
	Saida2012	50	1.38	20	2.8
Literature	Myhr1999	32	1.6		
	Filippi2006	548	1.73		
	Polman2006	315	1.3		
	Hauser2008	35	1.375		
	Kappos2008	65	1.125		
	Giovannoni2010	437	0.91		
	Comi2001			120	3.5
	Comi2008			102	3.9

GdE T1 lesions at 6 months; without (**A**) and with **B** baseline enrichment.

Application 1: no baseline enrichment

- 3 in-house (individual level data), 6 literature trials
- N=1936 historical patients (504 in-house, 1432 literature)
- Approach for prior τ
  - Use half-normal
    - Ensures positive mass around small heterogeneity
    - Weight in reasonnable range (not "heavy tailed")
  - Need to compare within- (σ) and between-trial variability (τ)
  - Very large heterogeneity: "$\tau \approx \sigma = \text{SD}(\log Y)$"
  - $\hat{\sigma} \approx 2$ from in-house data
  - Set median$(\pi(\tau)) = \hat{\sigma}$ (i.e. *HN*(scale=3))
  - Alternative: *HN*(scale=1), centered between substantial and large heterogeneity (Spiegelhalter et al. (2004))

Application 1: results



Prior	n^*
HN(scale=3)	25
HN(scale=1)	29

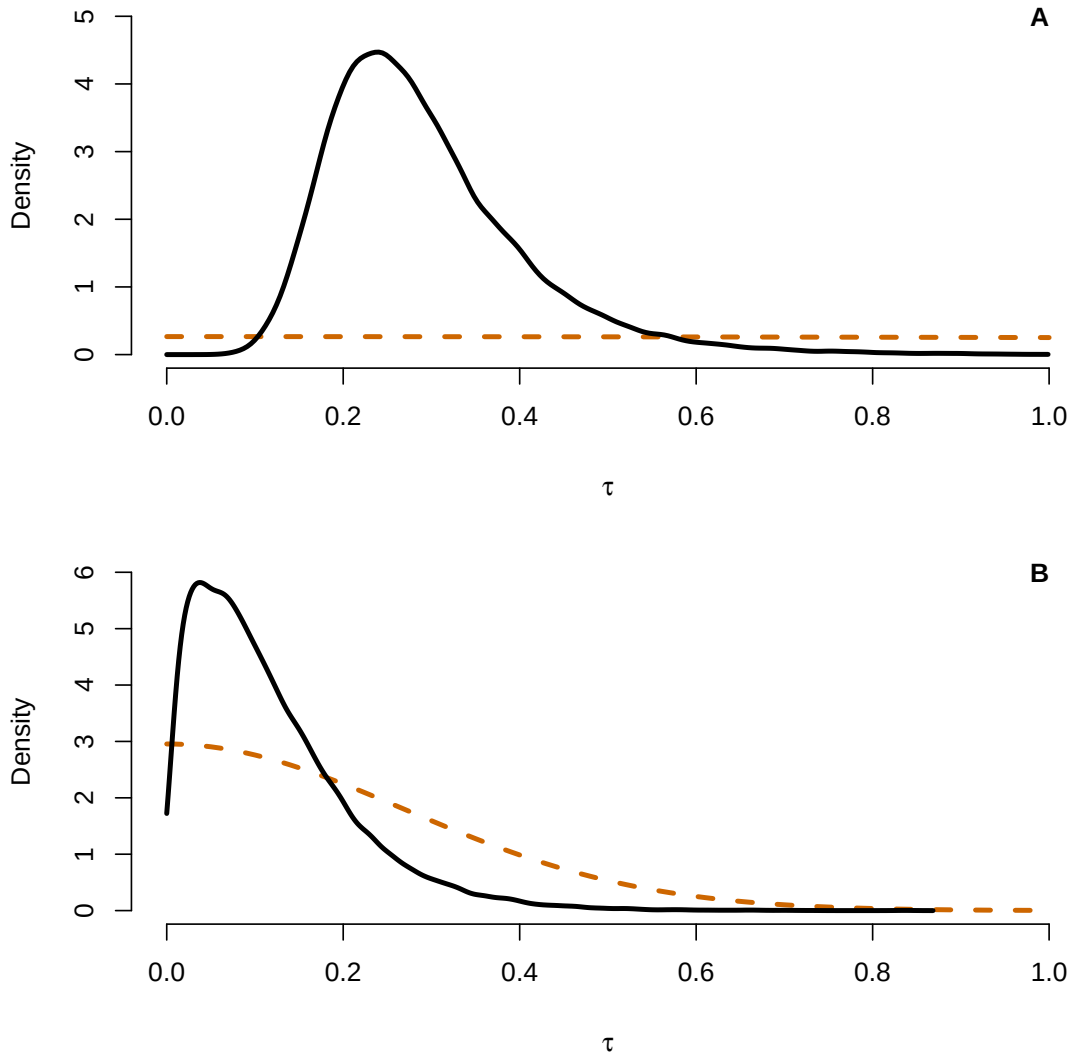
Heterogeneity	τ (fixed)	n^*
small	0.125	148
moderate	0.25	40
substantial	0.5	11
large	1	3
very large	2	1

Data indicate moderate to substantial heterogeneity.

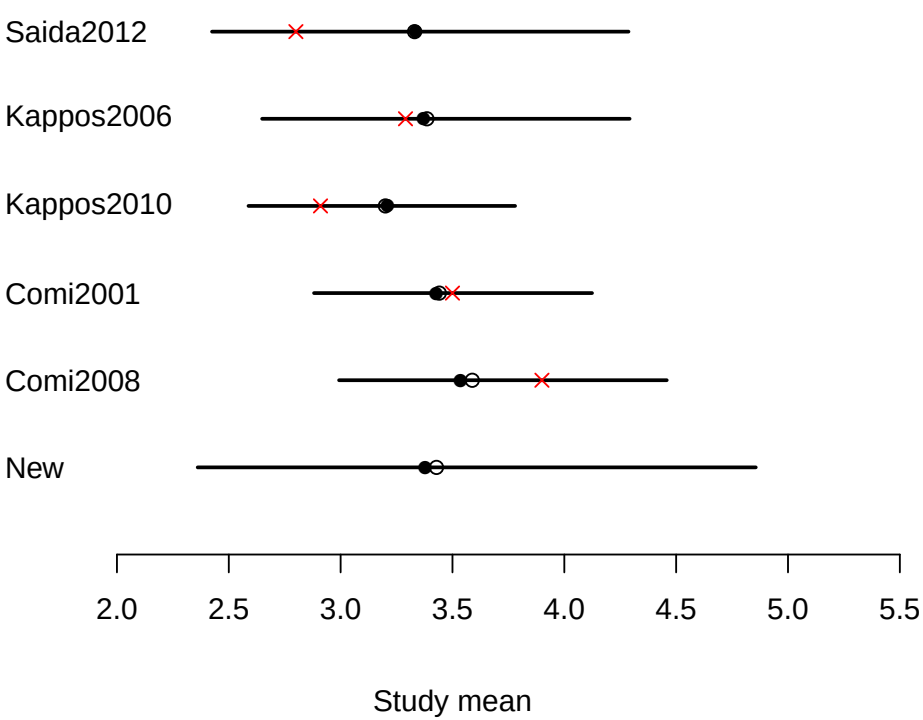
Application 2: use baseline enrichment

- Enrol patients with ≥ 1 lesion at baseline
- Minimal disease activity; homogeneity of study population
- Different trial than in application 1!
- 3 in-house (individual level data), 2 literature trials
- N=412 (190 in-house, 222 literature)
- Approach for prior $\tau \sim HN(\nu^2)$
  - Small number of trials, need informative prior
  - Less heterogeneity expected than in application 1
  - Set ν such that 90% quantile of prior is equal to 90% quantile of posterior for τ from application 1.

Priors and posteriors for τ



Application 2: results



Prior	n^*
HN(scale=0.26)	62

Heterogeneity	τ (fixed)	n^*
small	0.125	75
moderate	0.25	23
substantial	0.5	6
large	1	1
very large	2	0

Data indicate small to moderate heterogeneity.

Table of contents

Conclusions

- Historical data included in the analysis can lead to quicker and more ethical trials
  - Less patients on placebo
  - Potentially faster recruitement
- The actual number of patients to be replaced in new trial depends on various factors, not just n^*
  - Randomization
  - Objectives of trial; secondary and exploratory endpoints
  - Confidence of team with approach
- Analysis of new trial more difficult
  - Not for the Bayesian (”it’s just using a certain prior”)
  - ... but for the clinical team yes (not used to statistical analysis with different sources of information).

References

Spiegelhalter DJ, Abrams KR, Myles JP. (2004) Bayesian Approaches to Clinical Trials and Health-Care Evaluation. *Wiley*.
Berry et al. (2010) Bayesian Adaptive Methods for Clinical Trials. *Chapman & Hall*.
Pocock S. (1976) The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases*.
brahim and Chen (2000) Power prior distributions for regression models. *Statistical Science* 15:46-60.
Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter DJ. Summarizing historical information on controls in clinical trials, Clinical Trials. 2010.
Hobbs et al. (2011) Hierarchical Commensurate and Power Prior Models for Adaptive Incorporation of Historical Information in Clinical Trials. *Biometrics*.
Sormani et al. Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. J Neurol Sci. 1999.

Review of methods

See reviews in Spiegelhalter et al (2004), Berry et al (2010).

- Bias model (Pocock (1976)), $\theta_h = \theta + \delta_h$
- Power priors (Ibrahim& Chen (2000)),

$$\pi(\theta|D_0, \alpha) \propto L(\theta|D_0)^{\alpha}\pi_0(\theta),\ 0 \leq \alpha \leq 1$$

- Commensurate priors (Hobbs et al. (2011))
  - Explicitly parameterize similarity between D and D_0
  - Commensurability parameter τ
  - $\pi(\theta|D_0, \theta_0, \tau) \propto ...\pi(\theta|\theta_0, \tau)...$
- Random effects meta-analytic predictive (MAP) appraoch (Spiegelhalter et al (2004), Neuenschwander et al. (2010))

Structure of WinBUGS code

```
model{
      ## data generating model
      for(i in 1:ntot){ # exchangeable std log-means
        theta[i] ~ dnorm(eta, inv.tau2)
      ...
      }

      for(i in 1:nobs){ # likelihood: in-house data
            ...
        lambda[i] ~ dgamma(k.inv, kmu.inv[i])I(0.001,) # for num stability
        y[i] ~ dpois(lambda[i])
      }

      for(i in 1:nlit){ # likelihood: literature data
        ...
        xm[i] ~ dgamma(alpha[i], beta[i])I(0.001,)
        ym[i] ~ dpois(xm[i])
      }

      ## new study mean and new obs
      theta.new ~ dnorm(eta, inv.tau2)
      mu.new <- exp(theta.new)
      ...

      ## priors
      inv.tau2 <- 1/(tau*tau)
      tau ~ dnorm(0, pprec.tau)I(0,)
      ...
    }
```