Meta-Analytic Approaches to Using Historical Data in Clinical Trials

Bayes Pharma, 11 June 2014, London

Beat Neuenschwander, Head Statistical Methodology
Oncology, Biometrics and Data Management
Novartis Pharma AG, Basel

Joint work with:
Heinz Schmidli, Satrajit Roychoudhury (Novartis)
Sandro Gsteiger (University of Bern)
Tony O'Hagan (Sheffiled University)
David Spiegelhalter (Cambridge University)

NOVARTIS
ONCOLOGY

Outline

- Introduction with Examples
- Overview of Approaches
- Meta-Analytic Approaches
  - Meta-Combined and Meta-Analytic-Predictive Approach
  - Prior Effective Sample Size
  - Robustness
- More on Meta-Analytic-Predictive (MAP) Priors
- Conclusions

Introduction with Examples

1. Introduction and Examples

Informed Decision Making

- Informed decisions should be based on all relevant information
- In particular, when
  - information is sparse
  - new information is difficult to obtain
- Contextual or complementary data are often available

1. Introduction and Examples

Historical Data

- These data often referred too as «historical data»
  - But they may be come from a parallel experiment
  - Or, from data in the same experiment.
    E.g., in a clinical trial, from a similar subgroup
- Considering historical and current data is an example of evidence synthesis
- Various aspects to consider
  - methodological and practical issues and challenges
  - pros and cons

1. Introduction and Examples

Use of Historical Data: Pros & Cons

- Pros
  - **Design:** historical data are always used
    - This information puts the current experiment into perspective
    - For example: information about variability and expected effect sizsed drives sample size calculations
  - **Analysis:** historical data are rarely used.
    However, these data can improve the inference for key parameters
    - adjusted estimates (safeguard against extremes)
    - better precision

1. Introduction and Examples

Use of Historical Data: Pros & Cons

- Cons
  - What is relevant historical data?
    - Requires judgment about similarity of historical and current setting
    - Requires interaction between subject matter experts
  - How to incorporate historical data?
    - Requires a statistically principled approach
  - How much is the historical data worth?
  - What if historical data and actual data are in conflict?
    - Requires careful evaluation of the reasons
    - Problem can be mitigated by using a robust statistical approach

1. Introduction and Examples

Clinical Trials

- Use of historical data is attractive
  - Smaller sample sizes: e.g., smaller placebo group
  - More **ethical** (less placebo patients), or more **scientific** trials (learn more about new treatment)
  - Decreased costs and trial duration
- Historical data: various formats, e.g.
  - for control group only (our focus)
  - for effect parameter (mean difference, risk-ratio,...)
  - aggregate and/or individual data

1. Introduction and Examples

Novartis Experience

- Use of historical data
  - In all phase I oncology trials (to inform prior distributions)
  - In a substantial percentage of phase II trials
  - In special cases (e.g. non-inferiority trials)
- Experience overall positive
- However, there are challenges
  - Practical: drug development is highly regulated (company internal and external standards)
  - Practical: more time needed for study design
  - Methodological: innovative statistics

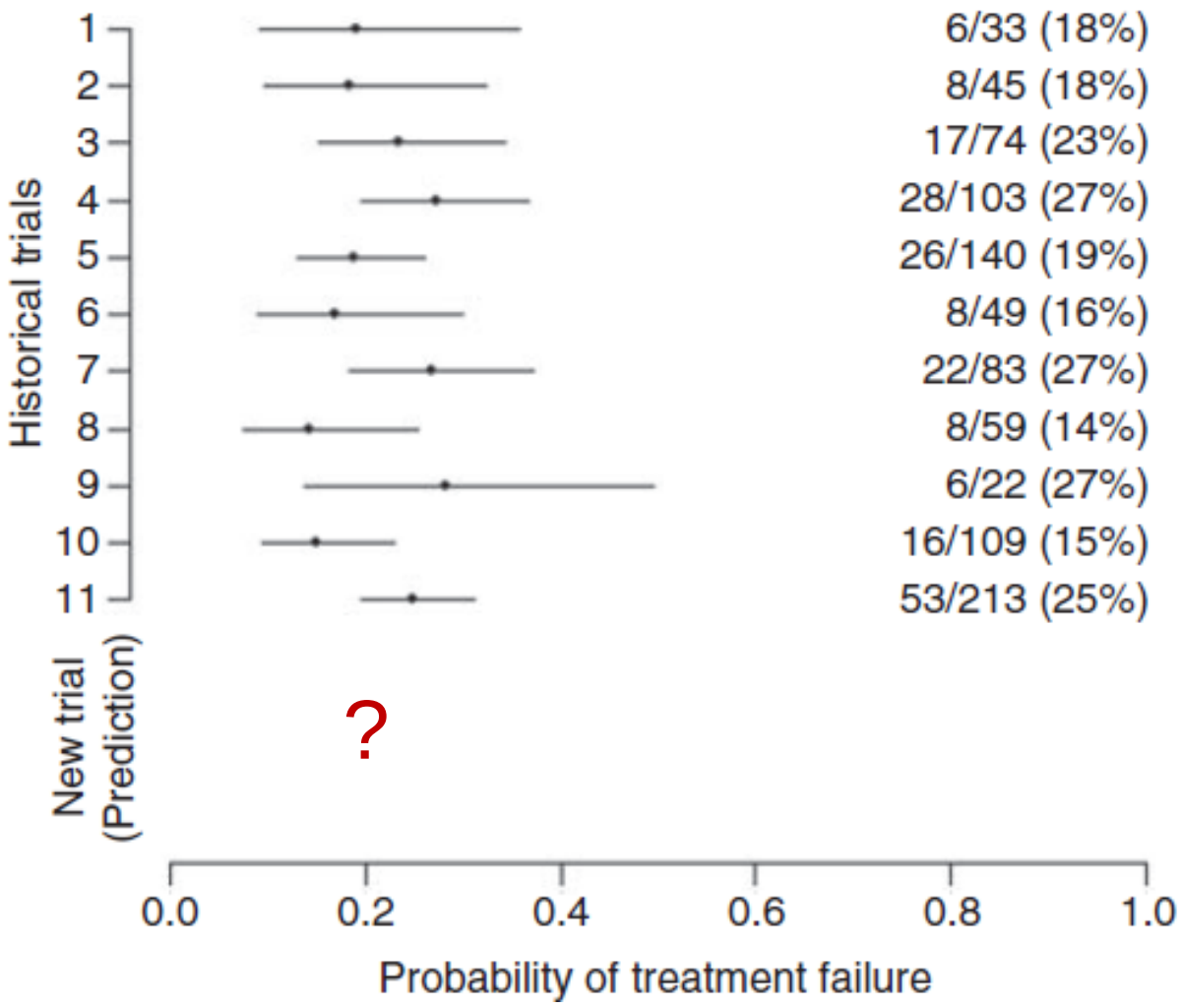
Example 1

Phase IV Trial

- Phase IV transplantation trial
- Binary outcome: treatment failure
- New treatment (T) vs. standard of care (C)
- Standard design: requires 450 patents per arm
- Historical data
  - 930 historical controls from 11 internal trials
  - Can these data be used to make control arm smaller?
  - See N et al. 2010

Example 1

Phase IV Trial: Control Data from 11 Historical Trials



NOVARTIS

Example 2

Phase II Design

- Phase II Trial in Ulcerative Colitis
- Outcome: clinical remission at week 8
- Placebo data from 4 external trials (363 historical controls) of similar design

Source	r/n	%
VanAssche (2007)	6/56	10.7
Feagan (2005)	9/63	14.3
Rutgeerts *et al.* I (2005)	18/121	14.9
Rutgeerts *et al.* II (2005)	7/123	5.7
Total	40/363	11.0

Example 3

Design of a Phase I Oncology Trial in Japan

- **Western** (on-going) first-in-human study
  - Objective: determine the maximum tolerated dose **(MTD)**
  - Endpoint: frequency of dose-limiting toxicity **(DLT)**
- Phase I study in **Japan** to find Japanese MTD
  - Often, no ethnic differences
  - For Japanese trial, can we make use of Western data?

Dose	100	200	400	800	1500	3000	TOTAL
# Patients	5	6	5	9	8	4	37
# DLT	0	0	0	0	1	3	4

↑ Tentative Western MTD

Overview of Approaches

2. Overview of Approaches

Find Relevant Historical Data

- 1st step: idenfity relevant historical data
  - Systematic Reviews methodology
    - E.g. Cochrane Handbook (Higgins and Green 2011)
  - Pocock's (1976) criteria
    - Inclusion/exclusion criteria for patient population
    - Type of study design
    - Exact definition of the outcome
    - Quality of study execution and management;
    - Potential biases due to time trends
  - Requires cross-functional expertise
    - A psychological barrier for many statisticians
    - May not lead to a unique set of trials ($\rightarrow$ sensitivity analyses)

2. Overview of Approaches

Basic Notation

- Index for
  - Historical data from *H* trials: ***1,...H***
  - Current/new trial: *****
- Data
  - Historical: $Y_1,...Y_H$ Current: Y_*
- Parameters
  - Historical: $\theta_1 , ... , \theta_H$ Current: θ_*
  - Use of historical data requires an assumption of similarity: formally expressed by parameter model for

$$\theta_1,..., \theta_H, \theta_*$$

2. Overview of Approaches

Approaches

- Original work in pre-clinical applications (1970s)
- The main approaches are
  1. Pocock's approach (bias model)
  2. Ibrahim & Chen Power Priors
  3. Meta-Analytic approaches (hierarchical models)
  - Approaches 1-3
    - Are conceptually and mathematically similar
    - Discount the historical data; see Spiegelhalter et al. 2004

2. Overview of Approaches

Pocock (1976)

- Differences between new and historical trial

$$\delta_h = \theta_* - \theta_h \quad (h=1,...,H)$$

- Assumption: no systematic biases
  This requires careful selection of historical data

$$\delta_h \sim N(0,\tau_\delta^2)$$

- The above model can be extended, but this requires additional assumptions
  - Bias assumptions $\rightarrow$ δ_h not centered at 0
  - Historical trials of different quality $\rightarrow$ different τ_δ
    - e.g., larger for observational, smaller for randomized controlled trials

2. Overview of Approaches

Power Priors (Ibrahim and Chen 2000)

- Prior for θ_*
  - For one historical trial:

$$p(\theta_*|Y_1) \propto L(\theta_*|Y_1)^a \times \pi_0(\theta_*)$$

  - Accounts for historical data via discounted likelihood
  - $a \in [0,1]$ determines the amount of discounting
    - $a = 1$: pooling of historical and new data; $a = 0$: no borrowing
  - Notes:
    - $\pi_0(\theta_*)$, a default non-informative prior
    - No formal model for θ_1 (historical) and θ_* (but see slides 20-21)
    - Extension to more trials: power parameters $a_1,...,a_H$

2. Overview of Approaches

Power Priors: Two Versions

- Fixed *a*
  - Discounting does not depend on how similar historical and new data are
- What about unknown *a* *(=α)?*
  - Prior on α ?

$$p(\theta_*, \alpha \mid Y_1) \propto L(\theta_* | Y_1)^{\alpha} \times \pi(\alpha) \times \pi_0(\theta_*)$$

  - This is not correct:
    - $L(\theta_* | Y_1)^{\alpha}$, conditional prior of θ_* given α; $\pi(\alpha)$ marginal prior of α
    - Normalizing constant on right-hand side depends on unknown α
    - Derivation of normalizing constant can be difficult
    - See Duan et al. 2006, N et al. 2009

2. Overview of Approaches

Power Priors: Unknown Power Parameter

- Simple Example: one trial with binary data
  - Uniform prior for power parameter α
  - Historical data: x_0 responders, y_0 non-responders, $n_0=x_0+y_0$
  - New data: x responders, y non-responders, $n=x+y$
  - Power priors
    - Original: $\propto \theta^{\alpha x0} (1-\theta)^{\alpha y0}$
    - Normalized: $= \Gamma(\alpha n_0+2)\, \Gamma^{-1}(\alpha x_0+1)\, \Gamma^{-1}(\alpha y_0+1)\, \theta^{\alpha x0} (1-\theta)^{\alpha y0}$
  - Data: historical $x_0/n_0=20/100$, new $x/n=20/100$
    - α posterior from original prior: $0.02\ (0.00,0.07)_{95\%}$???
    - α posterior from normalized prior: $0.57\ (0.07,0.98)_{95\%}$

2. Overview of Approaches

Hierarchical Modeling Approaches

- Data (within trials) suggests a hierarchical model that allows for between-trial heterogeneity
  - $\theta_1,\ldots,\theta_C,\theta_* \sim N(\mu,\tau^2)$
    - For normal-normal hierarchical model (see later slides), there is a 1-1 mapping between τ and *a* (power parameter)
    - Historical data: n observations with standard deviation σ (known)

$$a = \frac{1}{1 + 2n\tau^2/\sigma^2}$$

  - $\theta_1=\ldots=\theta_C=\mu,\ \theta_* \sim N(\mu,\tau^2)$
    - Commensurate prior approach (Hobbs et al. 2011,2013)
  - Note:
    - for one historical trial, the above approaches are equivalent

2. Overview of Approaches

1-1 Relationship: Power Parameter a vs. Between-Trial sd τ

- Example
  - Normal data (known standard deviation σ)
  - Hierarchical model (between trial sd τ)
- Power parameter *a* (%) as a function of
  - historical sample size *n* (one trial)
  - between trial-heterogeneity (σ^2/τ^2, see N et al 2010)

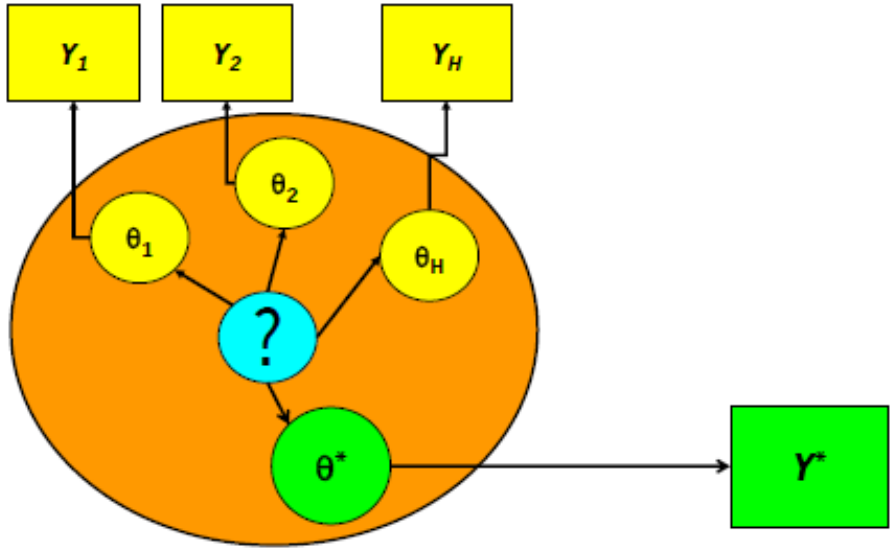
	large (4)	substantial (16)	moderate (64)	small (256)
n=25	7.0	20.0	60	80
n=50	4.0	10.0	40	70
n=100	2.0	7.0	20	60
n=250	0.8	3.0	10	30
n=500	0.4	2.0	6	20
n=1000	0.2	0.8	3	10

  - For moderate between-trial sd: historical data are worth
    - 20 subjects if n=100 (a=0.20),
    - 30 subjects if n=1000 (a=0.03)

Meta-Analytic Approaches

3. Meta-Analytic Approaches

Framework



- Meta-Analytic Approach
  - uses a data model $Y|\theta$, and a parameter model ?
  - infers the parameter of interest θ_*
    - at the end of the new trial (with Y_*),
    - or, at the design stage (without Y_*) $\rightarrow$ prior of θ_*

NOVARTIS

3. Meta-Analytic Approaches

Retrospective or Prospective Use of Historical Data

- Two MA approaches
  - ***Meta-Analytic-Combined (MAC)*** is **retrospective**
    - Perform a meta-analysis of historical data and current trial data
    - Parameter of interest: the parameter in the actual trial

$$\theta_* \mid Y_1,\ldots Y_H, Y_*$$

  - ***Meta-Analytic-Predictive*** **(MAP)** is **prospective**

1) At design stage of current trial:
   Perform MA of historical data data and obtain distribution of θ_*

MAP Prior: $\theta_* \mid Y_1,\ldots Y_H$

2) Combine MAP prior with current trial data Y_* (Bayesian analysis)

3. Meta-Analytic Approaches

MAC or MAP?

- ***Meta-Analytic-Combined (MAC)***
  - No prior for θ^* required at design stage
  - Only one analysis required, can be (non-)Bayesian
- ***Meta-Analytic-Predictive (MAP)***
  - Historical information about θ^* is explicitely stated at design stage
  - Historical data can then be ignored
  - Fully Bayesian analysis required
- ***MAC*** or ***MAP*** ? Which one is better?

3. Meta-Analytic Approaches

MAC and MAP Are Equivalent

- For a hierarchical model, MAC and MAP are equivalent
  - HM $\rightarrow$ data conditionally independent given parameters
  - That is: $Y_h \mid \theta_1,...,\theta_h,...\theta_H, \theta_* = Y_h \mid \theta_h$
  - Proof:

$$
\begin{aligned}
p(\theta_\star|Y_\star, Y_\mathcal{J}) &\propto p(\theta_\star, \theta_\mathcal{J}|Y_\star, Y_\mathcal{J}) \\
&\propto p(Y_\star, Y_\mathcal{J}|\theta_\star, \theta_\mathcal{J}) \times p(\theta_\star, \theta_\mathcal{J}) \\
&= p(Y_\star|\theta_\star) \times p(Y_\mathcal{J}|\theta_\mathcal{J}) \times p(\theta_\star, \theta_\mathcal{J}) \\
&\propto p(Y_\star|\theta_\star) \times p(\theta_\star, \theta_\mathcal{J}|Y_\mathcal{J}) \\
&\propto p(Y_\star|\theta_\star) \times p(\theta_\star|Y_\mathcal{J})
\end{aligned}
$$

3. Meta-Analytic Approaches

Normal-Normal Hierarchical Model (NNHM)

NNHM, very popular model

- Sampling model

$$Y_h \mid \theta_h \sim N(\theta_h, s_h^2) \qquad h = 1,...,H, *$$

- Parameter model

$$\theta_h \mid \mu,\tau \sim N(\mu, \tau^2) \qquad h = 1,...,H, *$$

- Inference: for θ_*
  - Challenge: what is τ? (in particular if H is small)
  - Classical: various ways to estimate τ
  - Bayesian: priors on μ (often flat) and τ (contextual)

3. Meta-Analytic Approaches

*Inference for **known** τ (with improper prior for μ)*

Basic formulas for fixed τ: Classical and Bayesian results are the same

Meta-analytic weights

$$w_h = 1/(s_h^2 + \tau^2)$$

Inference for μ

$$\hat{\mu} = \sum_h w_h Y_h / \sum w_h, \quad Var(\hat{\mu}) = 1/\sum w_h$$

Shrinkage factors

$$B_h = s_h^2/(s_h^2 + \tau^2)$$

Inference for θ_h

$$\hat{\theta}_h = B_h\hat{\mu} + (1 - B_h)Y_h, \quad Var(\hat{\theta}_h) = B_h(\tau^2 + B_h Var(\hat{\mu}))$$

Inference for new parameter $\theta^\star$

$$\hat{\theta}^\star = \hat{\mu}, \quad Var(\hat{\theta}^\star) = \tau^2 + Var(\hat{\mu})$$

Special casel: 1 historical trial: $\hat{\theta}^\star = Y_1, \quad Var(\hat{\theta}^\star) = s_1^2 + 2\tau^2$

3. Meta-Analytic Approaches

Unknown τ

- Discounting of historical data depends on τ
- For small number of trials
  - Classical
    - The various estimates can differ substantially
    - It is unclear how to adjust for estimation uncertainty
    - Proposal: for θ_*, t distribution with H-2 df (Higgins et al. 2009)
  - Bayesian
    - Conclusions can be sensitive to the prior
    - Judgment required about plausible values for τ

3. Meta-Analytic Approaches

τ-Priors. Spiegelhalter et al. 2004, Gelman 2006

- Various priors for τ
  - Uniform, inverse-sqrt-gamma, Half-Normal, Half-Cauchy...
  - Recommendation: use prior that puts
    - most of its mass to values that represent plausible heterogeneity
    - remaining probability to unanticipated heterogeneity (e.g. large)
  - Example: binary data, parameter = logit(p)
    - τ = 2 (1) correspond to very large (large) heterogeneity
    - Half-Normal priors (Spiegelhalter et al. 2004)

$\tau \sim$ Half-Normal(scale=1.0) $\rightarrow$ Pr(τ<2) $\approx$ 0.95

$\tau \sim$ Half-Normal(scale=0.5) $\rightarrow$ Pr(τ<1) $\approx$ 0.95

3. Meta-Analytic Approaches

Prior Effective Sample Size (ESS)

- Idea:
  - express prior as an equivalent number of subjects
  - the ***prior effective sample size (ESS)***
- What we know from conjugate analyses:
  - Binomial(n,p) data, Beta(a,b) prior
    - **Prior ESS: $n_0 = a+b$**
    - Posterior mean is a weighted average of prior mean and sample mean (with weights n_0 and n)
  - Similar results for normal, Poisson, exponential data, ...

3. Meta-Analytic Approaches
Approximating ESS

- More generally: ESS for MAP prior $\theta^* | Y_1,...,Y_H$
  - Approximate prior effective sample size n_*
    - Idea: sample sizes are (approximately) proportional to precisions
    - Under completeley homogeneous trials, $\tau = 0$

      $\Rightarrow$ $n_* = N = \Sigma_h n_h$ = total # of historical subjects

      $\Rightarrow$ $\mathrm{Var}_{\tau=0}(\theta_* | Y_1,..., Y_H)$ is proportional to 1/N
    - If $\tau>0$ (reality!) $\Rightarrow$ $\mathrm{Var}_{\tau>0}(\theta_* | Y_1,..., Y_H)$ is proportional to $1/n_*$

$$n_* = \frac{\mathrm{Var}_{\tau=0}(\theta_* \mid Y_1,...,Y_H)}{\mathrm{Var}_{\tau>0}(\theta_* \mid Y_1,...,Y_H)} \times N$$

    - More general approach to ESS, see Morita et al. (2008, 2012)

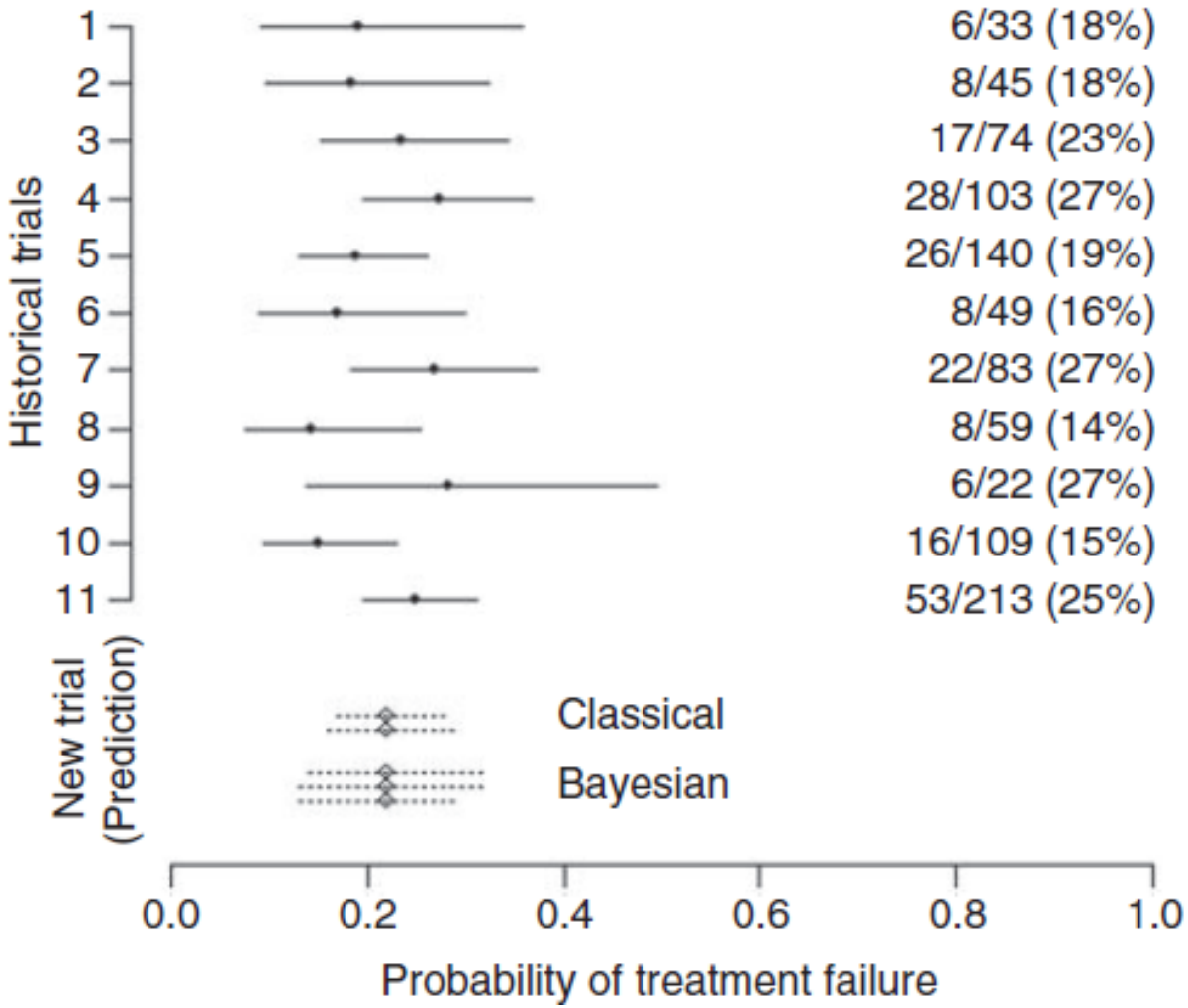
3. Meta-Analytic Approaches

Prior ESS for Example 1

- 11 historical trials with **N=930** patients
- Between-trial sd τ on log-odds scale **0.17 (0.01, 0.50)$_{95\%}$**
- 0.17: small/moderate

Results for log-odds θ_*

- Pooled: –1.27 (0.080)
- MAP: –1.29 (0.253)
- **Prior ESS**
  $n^* = 930 \times (0.08/0.253)^2$
  = 93



NOVARTIS

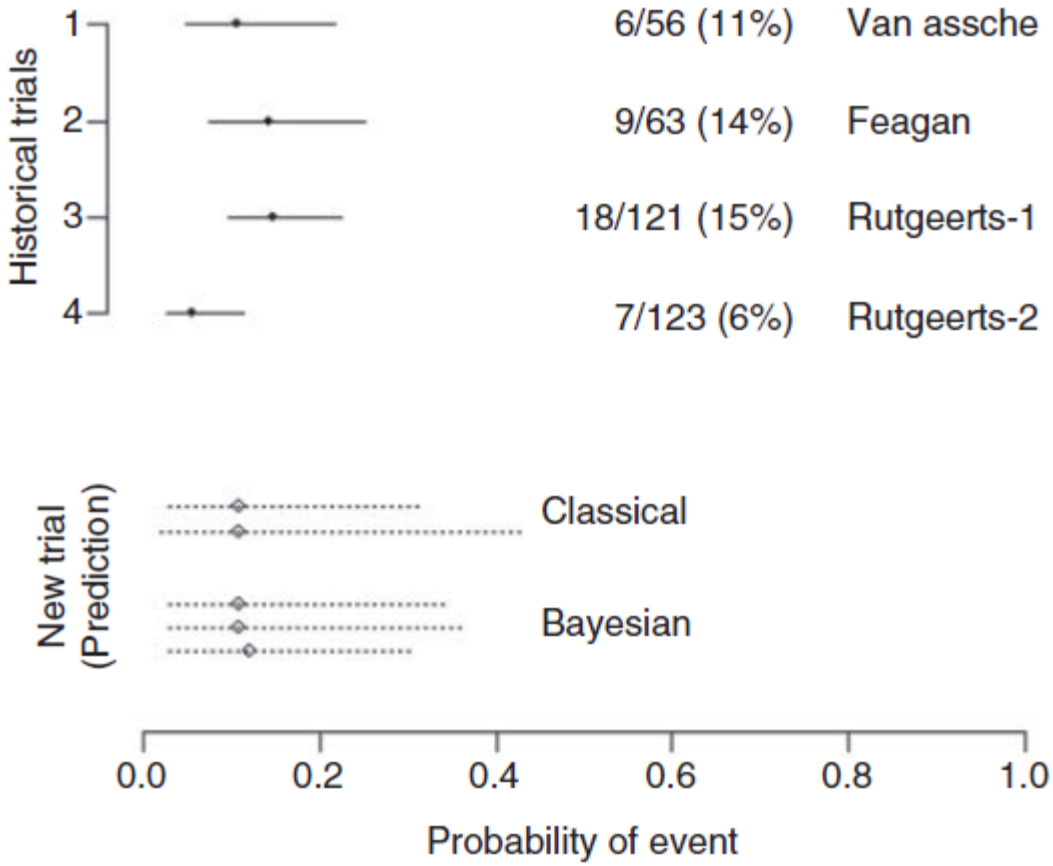
3. Meta-Analytic Approaches

Prior ESS for Example 2

- 4 historical trials with **N=363** patients
- Between-trial sd τ on log-odds scale **0.41 (0.03, 1.39)$_{95\%}$**
- 0.41: substantial

Results for log-odds θ_*

- Pooled: –2.01 (0.169)
- MAP: –2.08 (0.690)
- **Prior ESS**
  n* = 363×(0.169/0.690)^2
  = 22

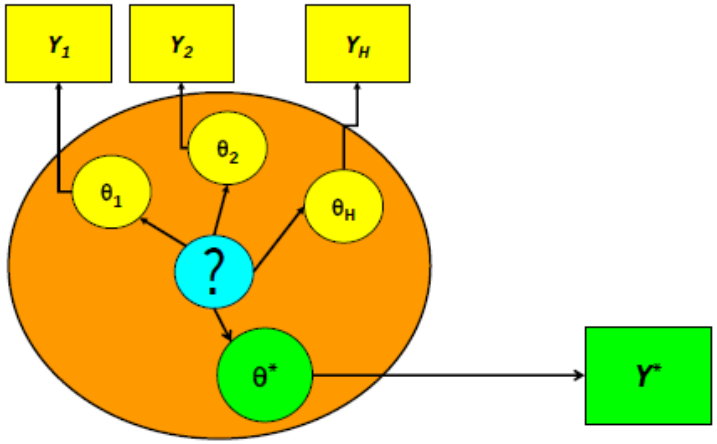


NOVARTIS

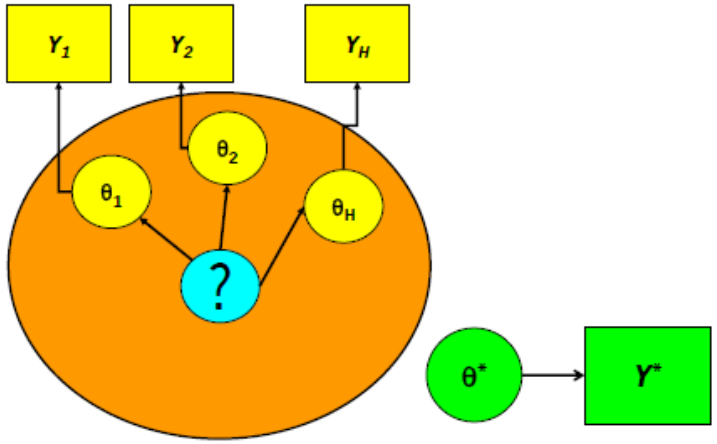
3. Meta-Analytic Approaches

Robust Meta-Analytic Priors

Similarity Scenario ($\rightarrow$ MAP prior)



Dissimilarity Scenario



- Conflict between historical data and actual data
  - Similarity of parameters is violated
  - Solution: robust priors (O'Hagan 1979); heavy-tailed (t or mixture)
- Robustifed MAP prior

  w $\times$ (MAP-prior) + (1-w) $\times$ (weakly-informative prior)

NOVARTIS

3. Meta-Analytic Approaches

Example 3: Robust MAP Priors

- **Western** (on-going) first-in-human study
  - Objective: determine the maximum tolerated dose (MTD)
  - Endpoint: frequency of dose-limiting toxicity (DLT)
- Phase I study in **Japan** to find Japanese MTD
  - Often, no ethnic differences
  - For Japanese trial, can we use of Western data?

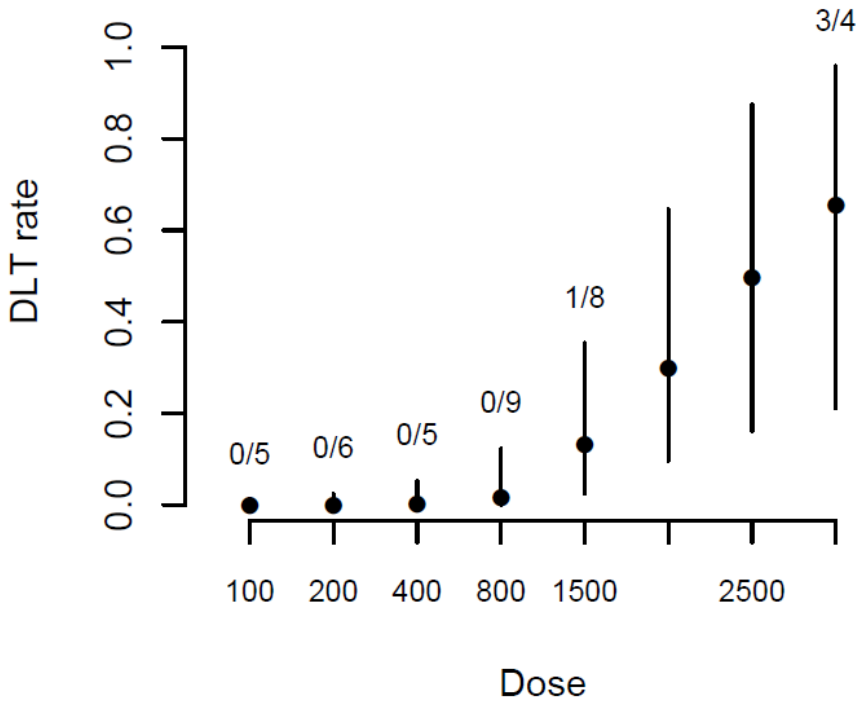
Dose	100	200	400	800	1500	3000	TOTAL
# Patients	5	6	5	9	8	4	37
# DLT	0	0	0	0	1	3	4

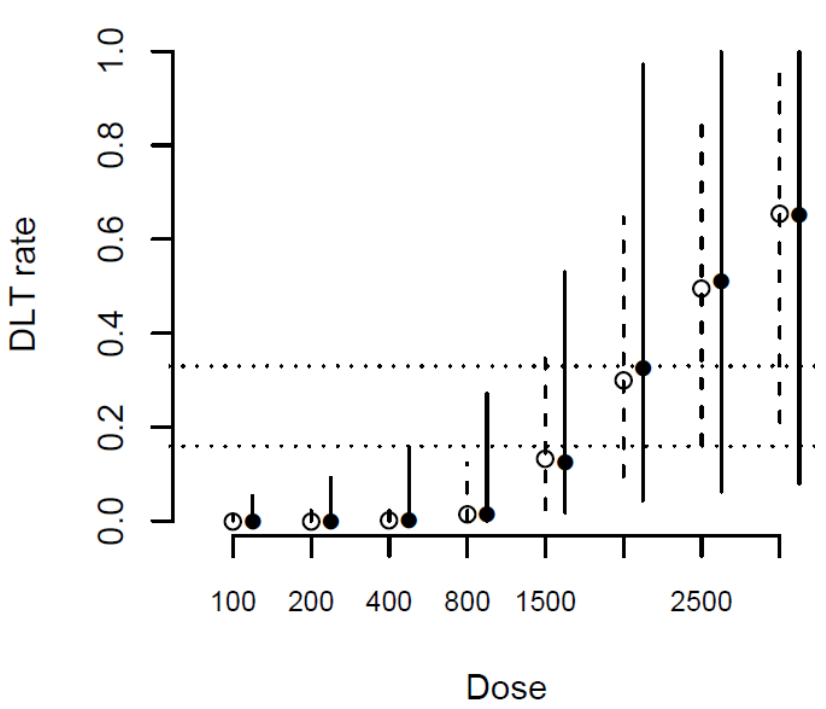
↑ Tentative MTD

3. Meta-Analytic Approaches

Example 3: MAP Prior for Similarity Scenario

- Model: logistic regression, with bivariate-normal prior for (α,β)
- Left: posterior from Western data
- Right: posterior from Western data (dotted line),
  MAP prior for Japan (solid line), under substantial heterogeneity





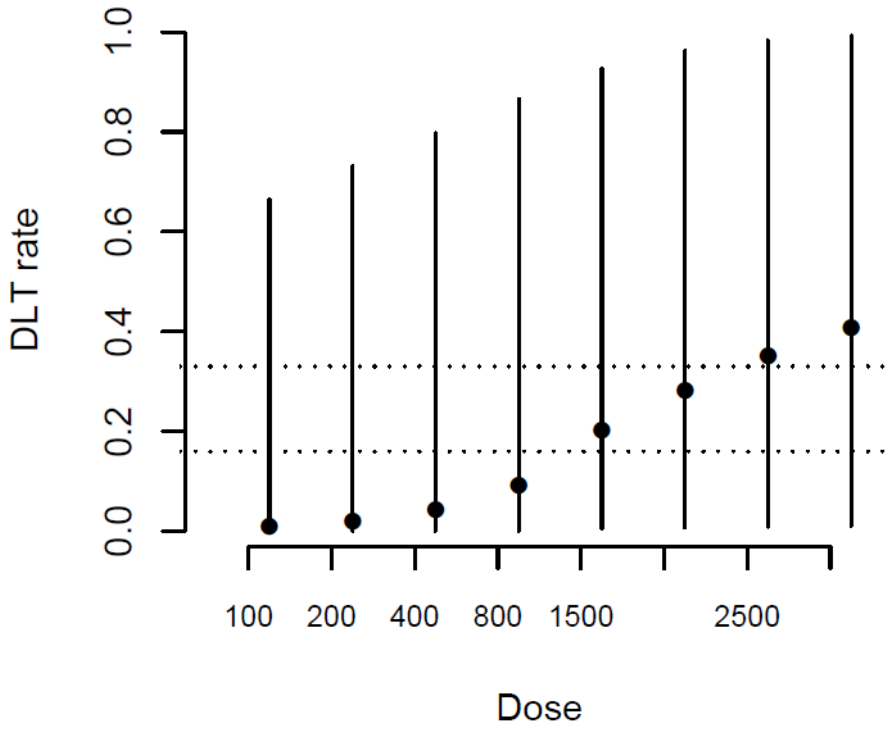
NOVARTIS

3. Meta-Analytic Approaches

Example 3: Weakly-Inf Prior for Dissimilarity Scenario

But what if ...

- There are relevant ethnic differences
- Better: to use weakly-informative prior (Figure)



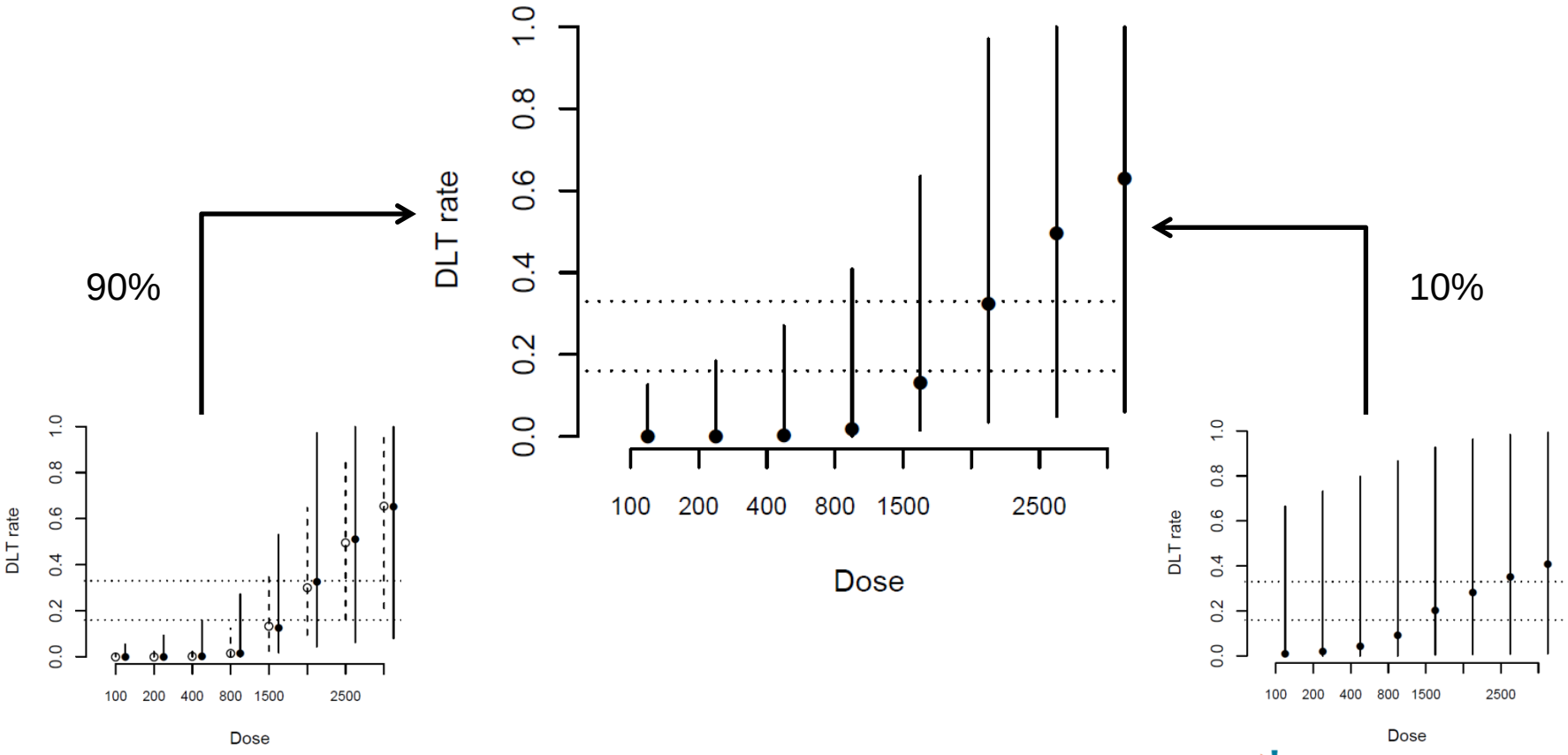
NOVARTIS

3. Meta-Analytic Approaches

Example 3: Robustification (Mixture Prior)

Mixture prior for the two scenarios, with the weights

- 90% for similarity scenario, 10% for dissimilarity scenario



NOVARTIS

3. Meta-Analytic Approaches

Example 3: Two Data Scenarios

- Design properties
  - Assess operating characteristics
  - Assess data scenarios that may arise in the trial

Dose	100	200	400	800	1200	1500	3000
Western Data							
#DLT/#Pts	0 / 5	0 / 6	0 / 5	0 / 9		1 / 8	3 / 4
Japan: scenario 1 (similarity)							
			0 / 3	0 / 3	0 / 3	1 / 3	
Japan: scenario 2 (dissimilarity)							
			0 / 3	2 / 3			

3. Meta-Analytic Approaches

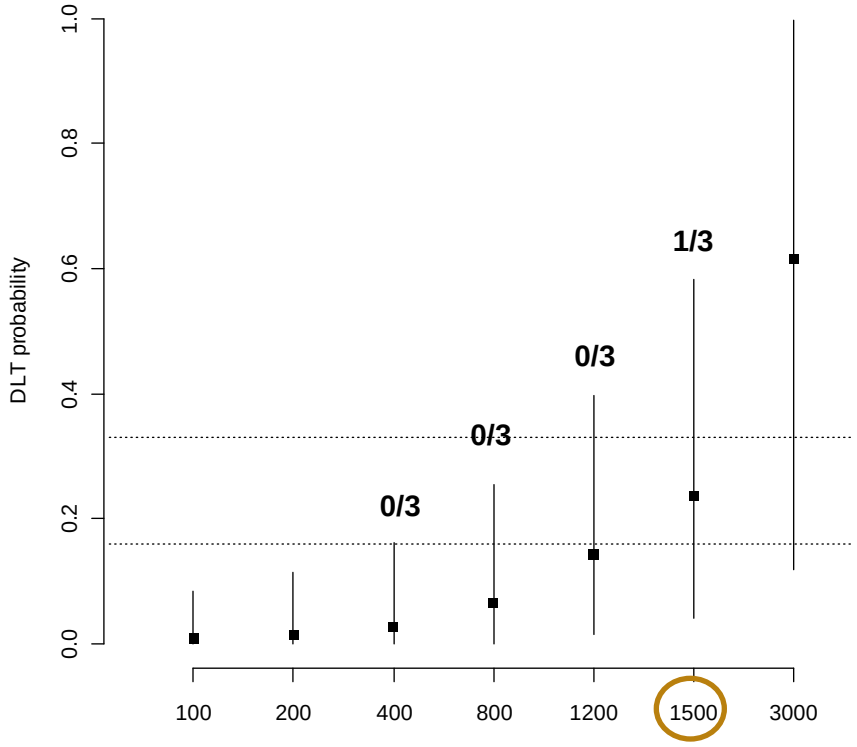
Example 3: Posteriors for Two Data Scenarios

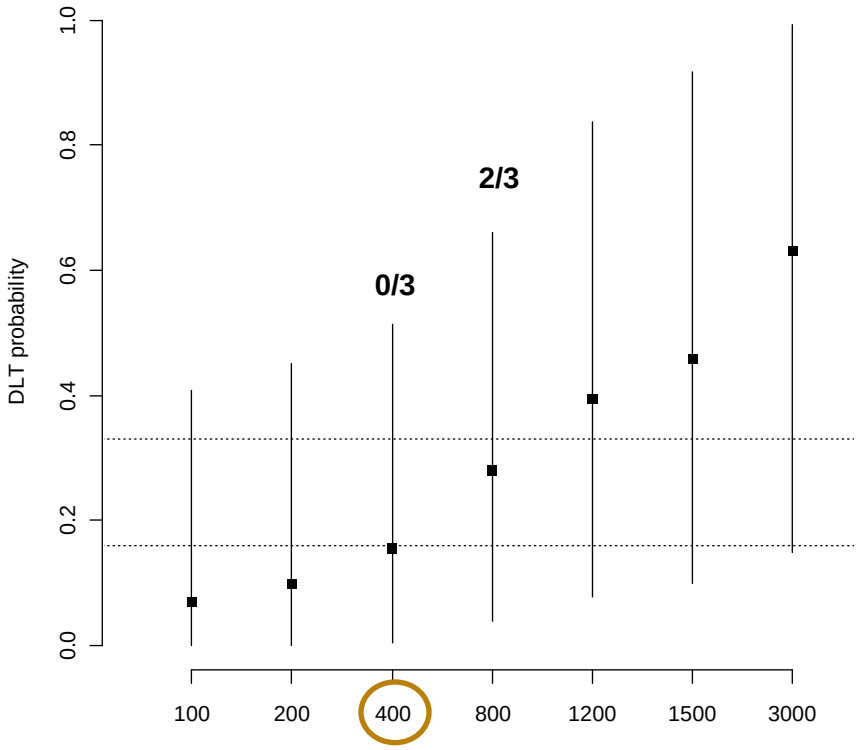
Similarity scenario

- Less uncertainty compared to prior
- Recommendation: **retest at 1500**
- Good borrowing from Western data

Dissimilarity scenario

- **More uncertainty** compared to prior
- Recommendation: **de-escalate to 400**
- Good robustness



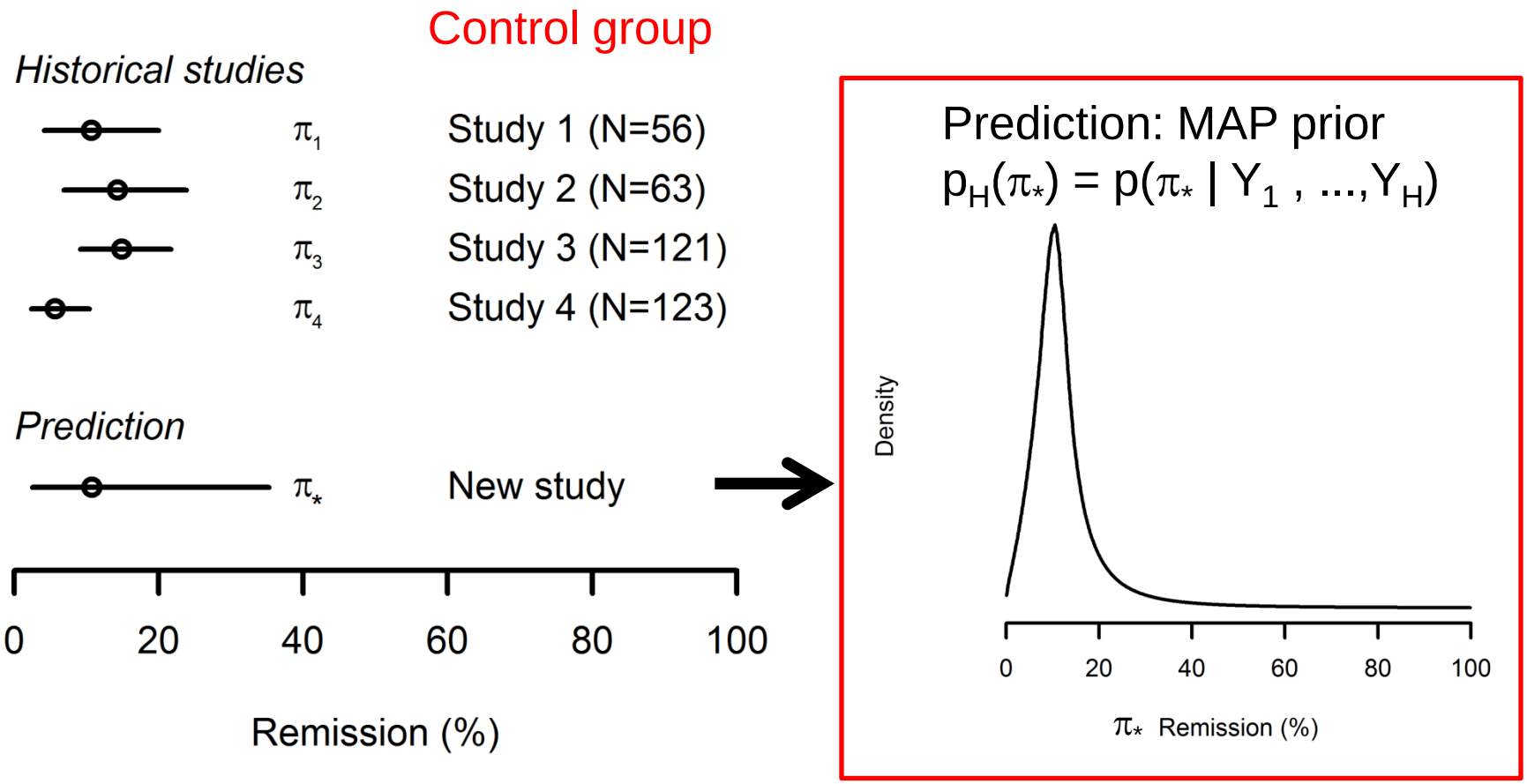


NOVARTIS

More on MAP Priors

4. More on MAP Priors

Example 2 Revisited



(density plot from MCMC sample)

NOVARTIS

4. More on MAP Priors

Approximating the MAP Prior

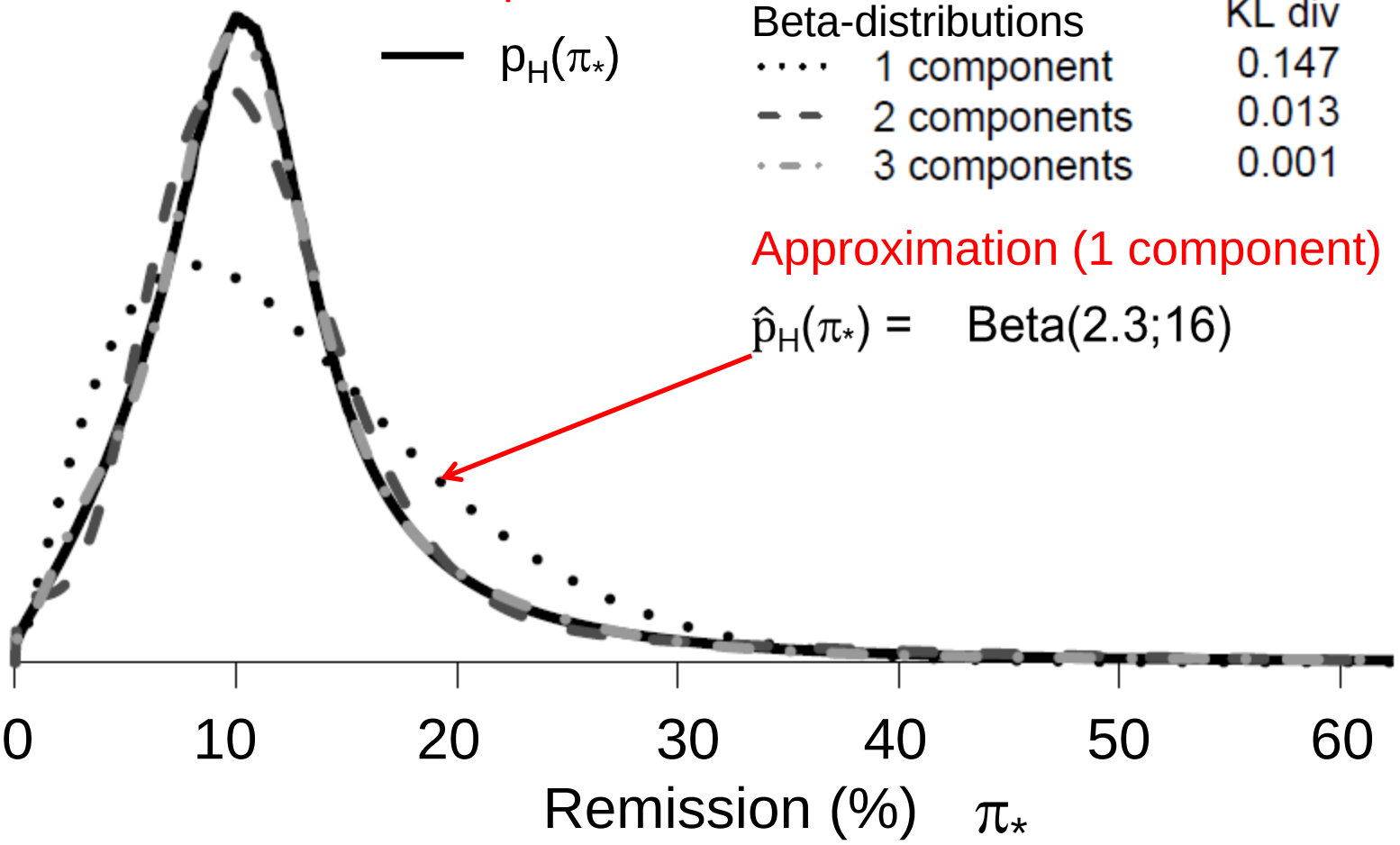
- MAP prior
  - Not available analytically (just MCMC sample), but can be well approximated by mixture of conjugate priors

    Dallal and Hall (1983), Diaconis and Ylvisaker (1985)

  - Mixture of conjugate priors. Advantages
    - *Easy communication*:
      discussions with clinical trial team, health authorities, ethics commitees, study protocols, publications
    - A*nalytical posterior calculation*
      $\rightarrow$ fast operating characteristics

4. More on MAP Priors

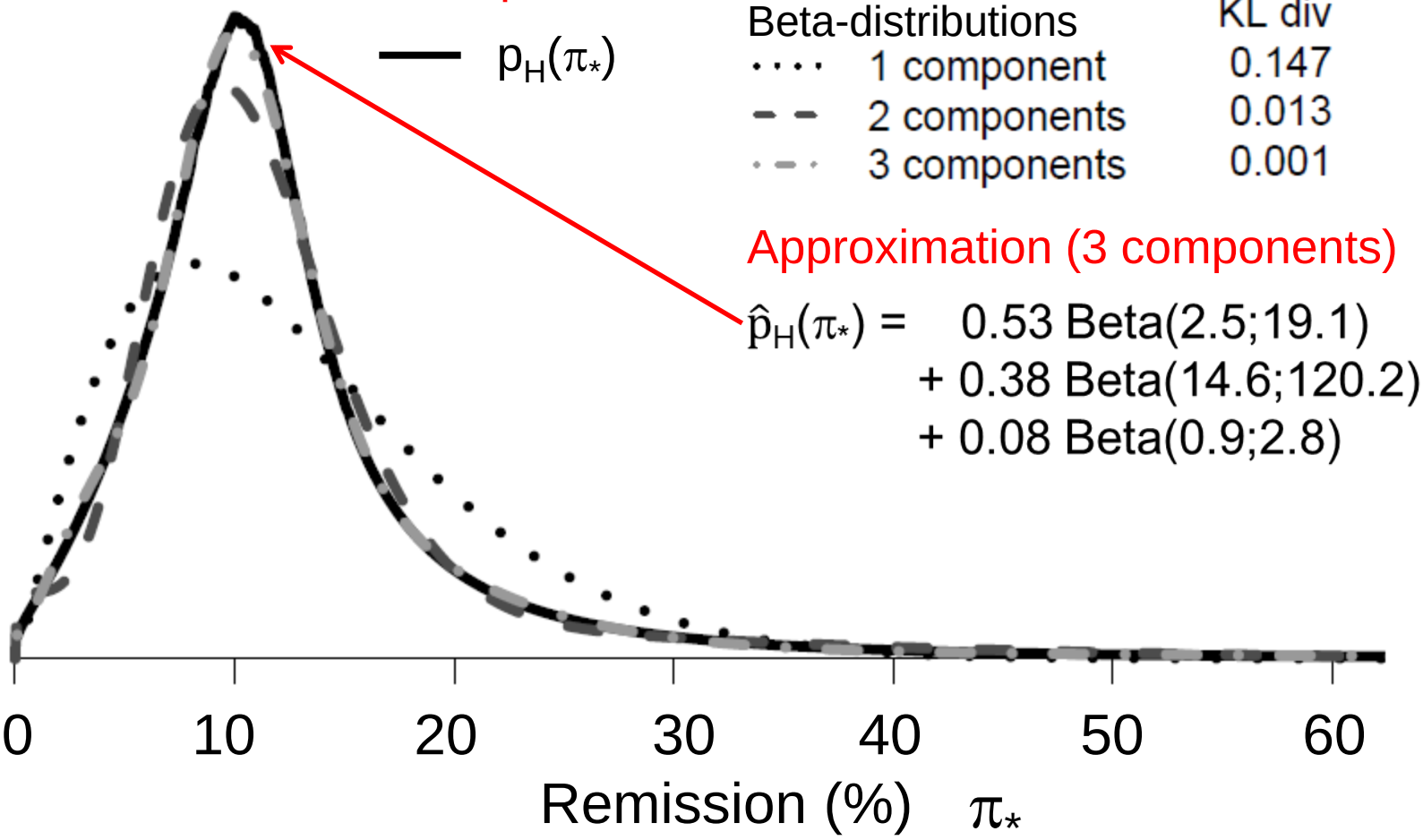
Example 2: MAP Prior approximated by single Beta



NOVARTIS

4. More on MAP Priors

Example 2: MAP Prior approximated by 3-comp Beta Mixture



NOVARTIS

4. More on MAP Priors

Robustness

- Prior-data conflict
  - *Conjugate priors*: fixed prior-data compromise
  - *Heavy-tailed priors* : prior discarded under conflict

O'Hagan (1979), O'Hagan and Pericchi (2012)

- MAP priors
  - Typically heavy-tailed, hence naturally robust
  - Further robustness and more rapid adaptation to prior-data conflict by adding weakly-informative component:

$$w \times \text{MAP} + (1\text{-}w) \times \text{Uniform} \quad \text{e.g. } w = 0.9 \text{ or } 0.5$$

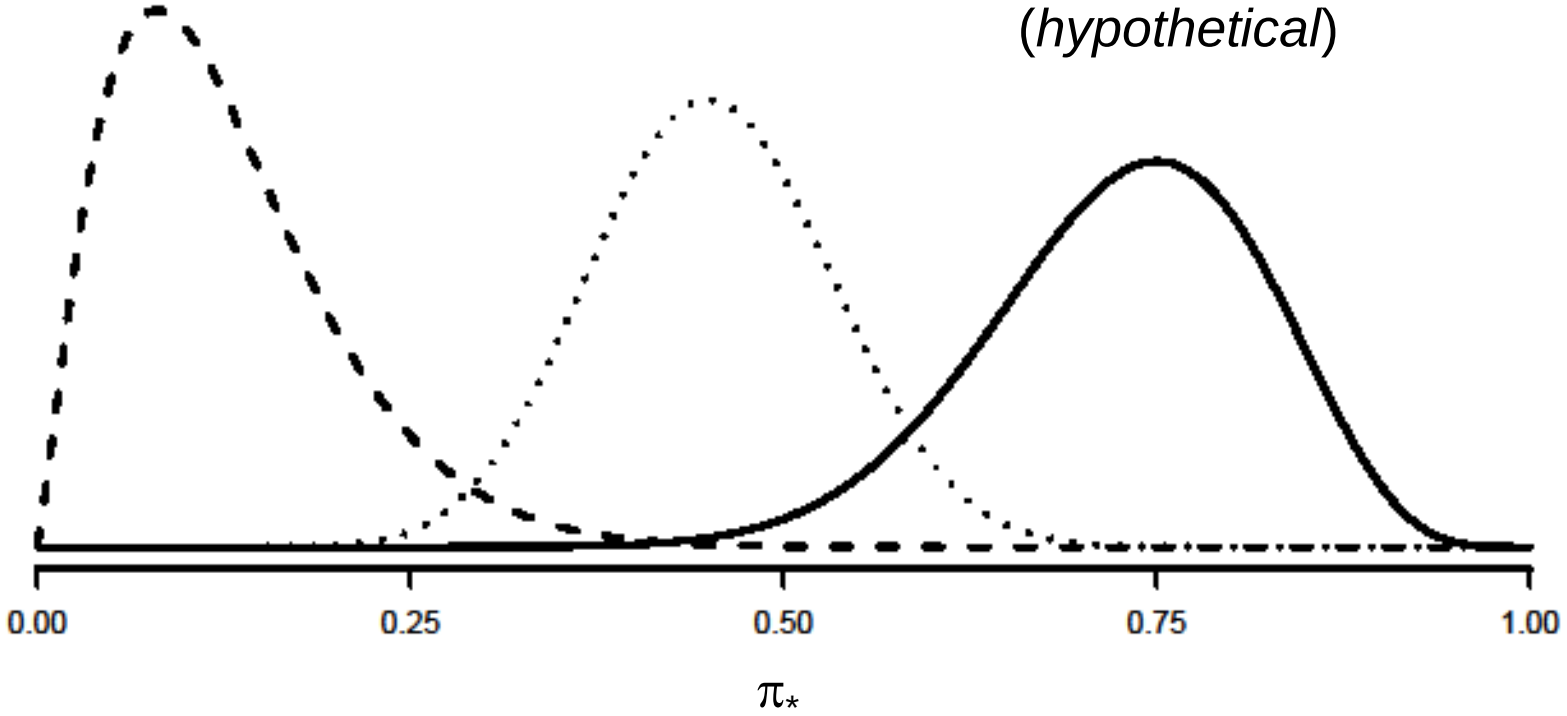
4. More on MAP Priors

Non-Robustness of Conjugate Prior

Conjugate prior
Beta(2.3,16)

Posterior
Beta(17.3,21)

Likelihood
15 / 20
(*hypothetical*)

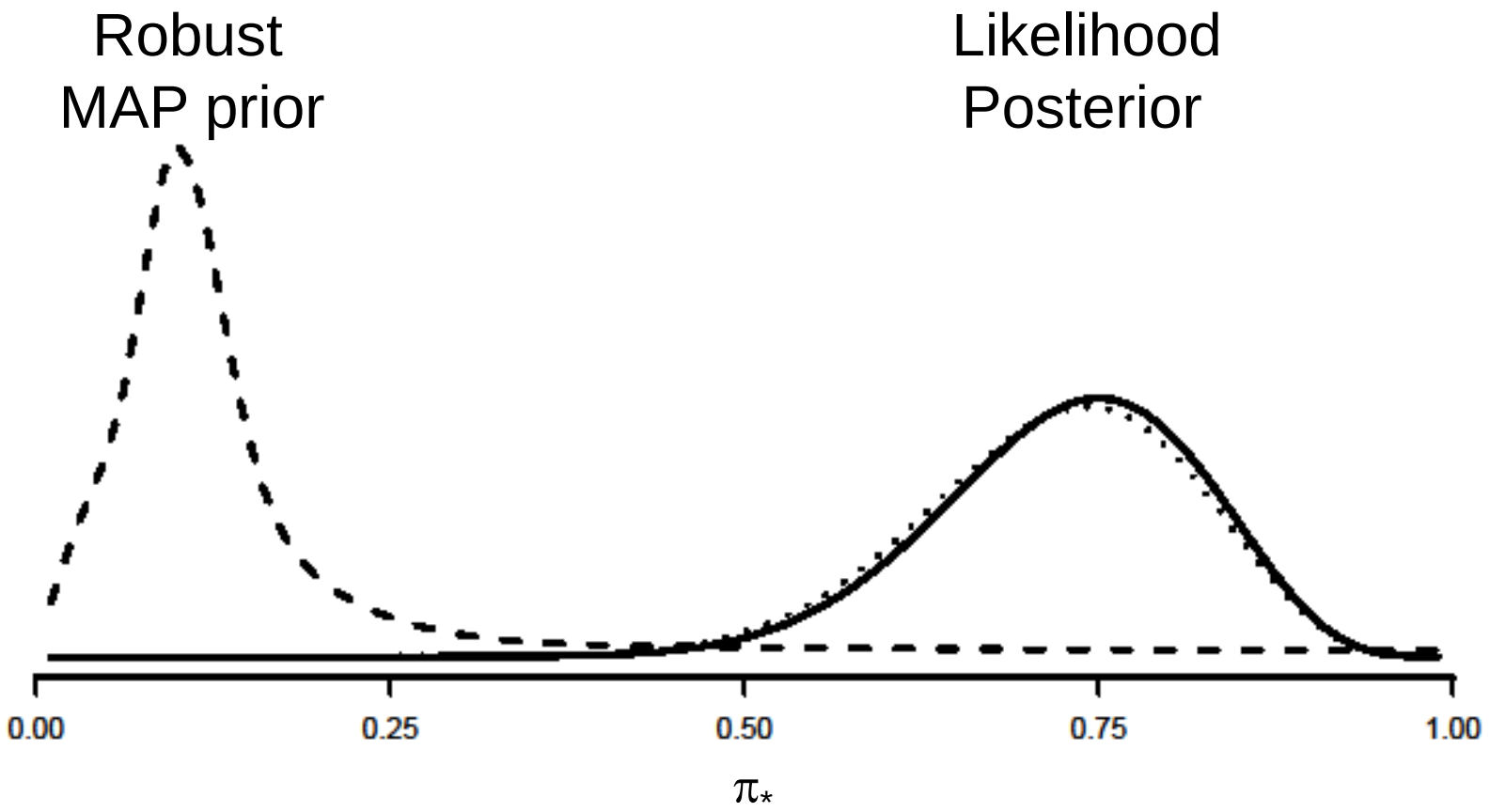


"Bayesian - One who, vaguely expecting a horse and catching a glimpse of a donkey, strongly concludes he has seen a mule". Stephen Senn

NOVARTIS

4. More on MAP Priors

Robustness of MAP Prior

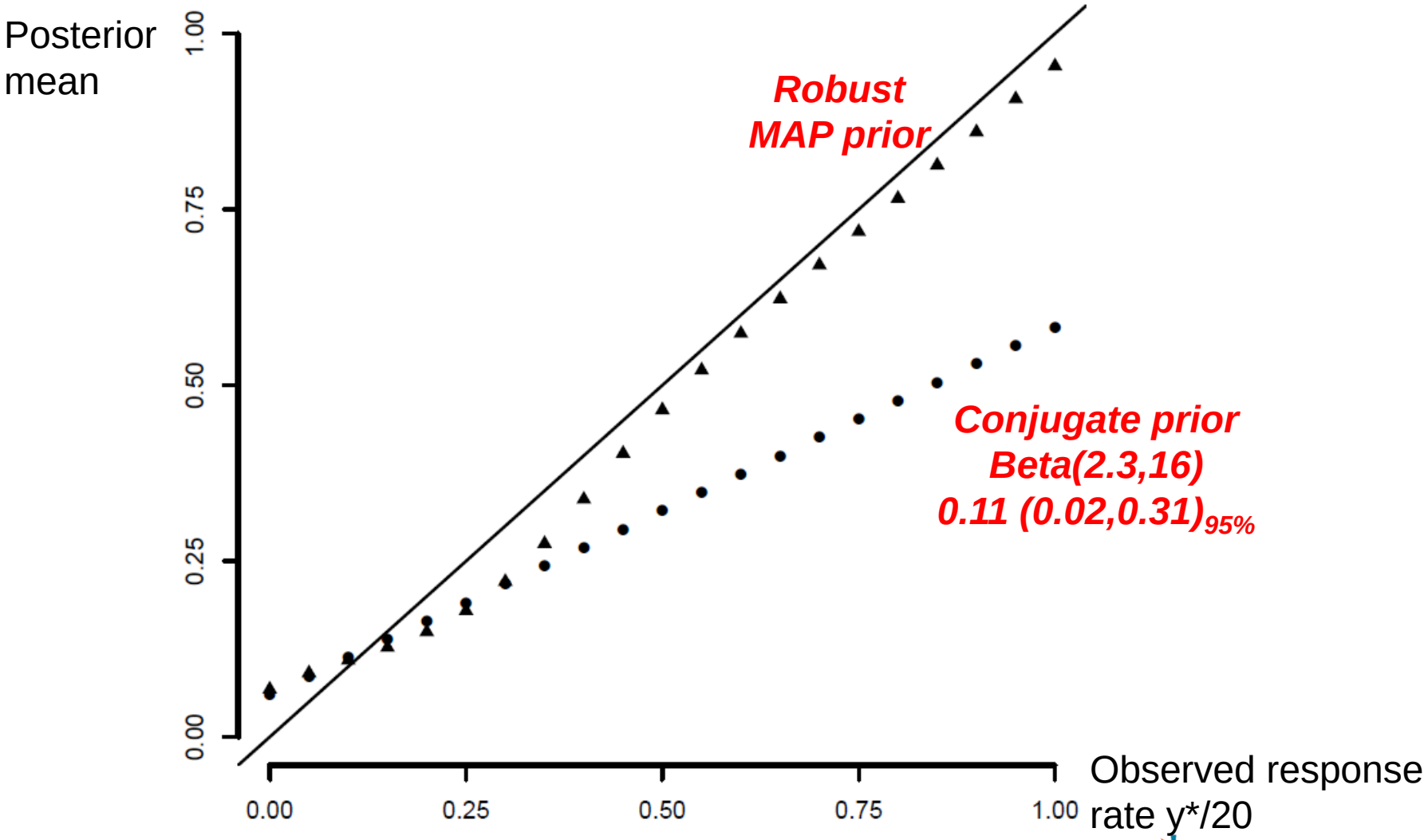


- MAP = 0.53 Beta(2.5,19.1)+0.38 Beta(14.6,120.2)+0.08 Beta(0.9,2.8)
- Robust MAP = 0.9 x MAP + 0.1 x Beta(1,1)

NOVARTIS

4. More on MAP Priors

Estimates for Simple Conjugate and Robust MAP Prior



NOVARTIS

4. More on MAP Priors

Operating Characteristics (OC): Summary

- Frequentist properties (OC) for robust MAP priors
  - *Estimation*:
    - Bias well-controlled
    - MSE: better for MAP priors compared to weakly-informative priors if prior is well-specified
  - Testing
    - Success criterion = 1-α posterior probability for $\delta = \theta_T - \theta_* > 0$
    - Type-I error: some inflation (or deflation), but fairly well controlled
    - Power: gain in power compared to weakly-informative prior

4. More on MAP Priors

Operating Characteristics (Estimation): Two Designs

- Compare Control vs. Test
  - Control vs. treatment effect: $\delta = \theta_T - \theta_*$
  - Control prior worth n^* patients:

$$\theta_* \sim N(\theta_0, \sigma_0^2), \qquad \sigma_0^2 = \sigma^2 / n^*$$

  - Assume no information for test treatment (flat prior for θ_T)
  - Two Designs
    - Standard **B**alanced Design (B), with sample sizes *n*
    - **H**istorical Data Design (H): save n^* control patients

	C	C-prior	T
B: Balanced Design	*n*	-	*n*
H: Historical Data Design	*n*-n^*	n^*	*n*

4. More on MAP Priors

Operating Characteristics (Estimation): MSE

- Mean-squared error (MSE) for mean difference δ

$$\mathrm{MSE}_{(H)} > \mathrm{MSE}_{(B)} \quad \Leftrightarrow \quad |\theta - \theta_0|/\sigma_0 > 1$$

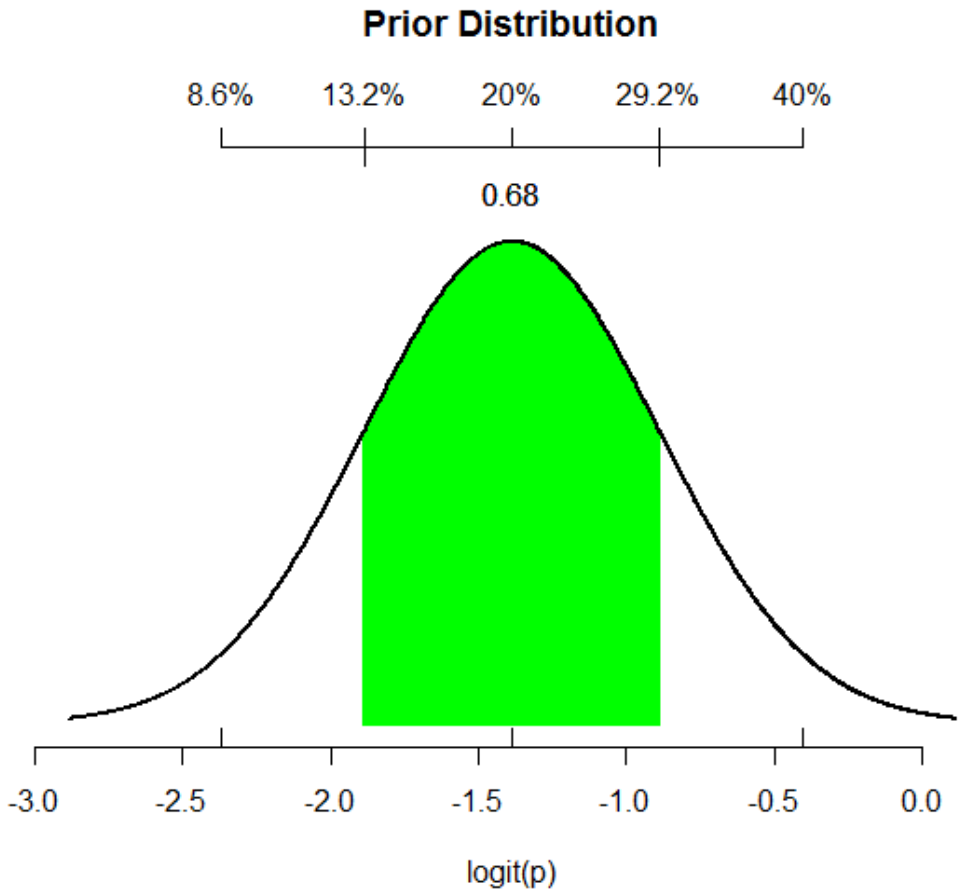
- **H**istorical data design better than **B**alanced design
  - if true parameter is less than one standard deviation away from the prior mean
  - i.e., if true parameter is in the 68% interval of the prior
- There is a benefit if prior is well-specified

4. More on MAP Priors

Operating Characteristics (Estimation): MSE - Example

Example: Binary data

- Control response rate
- Prior:
  - mean = 0.2, weight $n_* = 25$
- Normal approximation
  - logit(0.2) = –1.386
  - $(1/p+1/(1-p)) / n_* = 0.5^2$
- Prior:
  - logit(p) ~ $N(-1.386, 0.5^2)$
  - 95%-interval: ***0.086 to 0.4***



- For MSE, H-design better than B-design if $p \in (0.13, 0.29)$

NOVARTIS

4. More on MAP Priors

Operating Characteristics (Testing): Comparison of Priors

- Test treatment vs. Control, binary endpoint
  - Vague prior for test treatment: Beta(1,1)
  - Informative prior for control, e.g.

    ESS

    i. Beta: simple conjugate Beta(4,16) prior: 0.19 $(0.06,0.40)_{95\%}$

    ii. Mix90: $0.9 \times$ Beta(4,6) $+ 0.1 \times$ Uniform

    iii. Mix50: $0.5 \times$ Beta(4,16) $+ 0.5 \times$ Uniform

    iv. Unif: Uniform prior
  - Robust prior on control discarded in case of prior-data conflict – may lead to inconclusive results
  - An adaptive design can reduce this risk (Hobbs et al. 2013)

4. More on MAP Priors

Operating Characteristics (Testing): Adaptive Design

- Two-stage adaptive design
  - Target sample size at end of trial:
    - n = 40 for control, m = 40 for test
  - Stage 1:
    - n_1 = 15 for control
    - m_1 = 20 for test
  - Interim analysis: for control, get interim ESS_C
  - Stage 2 of adaptive design:
    - 40 - ESS_C for control
    - 20 for test

4. More on MAP Priors

Operating Characteristics (Testing): Type-I Error, Power

Control Rate	Mix50	Mix90	Beta	Unif	Expected Sample Size (Control Group) Mix50	Mix90	Beta	Unif
Type-I Error (δ =0)								
0.1	0.6	0.1	0.0	1.8	28	20	20	40
0.2	2.5	1.5	1.6	2.3	26	20	20	40
0.3	3.9	5.5	6.1	2.4	29	21	20	40
0.5	**3.4**	**12.3**	**26.0**	**2.8**	**37**	**27**	**20**	**40**
Power (δ =0.3)								
0.1	92	81	82	90	28	20	20	40
0.2	88	86	88	82	26	20	20	40
0.3	83	88	93	80	29	21	20	40
0.5	**78**	**85**	**99**	**82**	**37**	**27**	**20**	**40**

Schmidli et al. (2014, submitted)

Conclusions

5. Conclusions

- Use of historical (trial-external) data is
  - attractive
  - ambitious
  - ambiguous
- *Attractive*
  - more information should lead to better inference, and, subsequently, to better decisions
  - various potential benefits: smaller control groups, more ethical trials, cost savings

5. Conclusions

- *Ambitious*
  - Requires upfront work: find relevant data
  - Statistically more challenging
  - MA approaches (various dialects) are useful
  - Robust approaches look promising
- *Ambiguous*
  - Compromise between acceptable frequentist and Bayesian metrics is needed
- Clinical trials: the topic is important, and its importance will most likely grow in the near future

References

References

- Baeten D et al. (2013). Anti-interleukin-17A monoclonal antibody secukinumab in ankylosing spondylitis: a randomized, double-blind, placebo-controlled trial. *The Lancet* 382, 1705-1713.
- Bernardo JM, Smith AFM (1994). *Bayesian Theory*. Chichester: Wiley.
- Berry DA (2006) Bayesian clinical trials. *Nature Review Drug Discovery*, 5:27-36.
- Berry, Carlin, Lee, Müller (2010. *Bayesian Adaptive Methods for Clinical Trials*
- **Chen and Ibrahim (2006).** The Relationship Between the Power Prior and Hierarchical Models. Bayesian Analysis
- **Dallal S, Hall W (1983).** Approximating priors by mixtures of natural conjugate priors. *Journal of the Royal Statistical Society, Series B* 45, 278-286.
- Dempster A, Selwyn M, Weeks B. (1983). Combining historical and randomized controls for assessing trends in proportions. *Journal of the American Statistical Association* 78,221-227.
- **Diaconis P, Ylvisaker D (1985).** Quantifying prior opinion, In *Bayesian Statistics* 2, Bernardo et al. (eds), 133-156. The Netherlands: Elsevier.
- DiScala L, Kerman J, Neuenschwander B (2013). Collection, synthesis, and interpretation of evidence: a proof-of-concept study in COPD. *Statistics in Medicine*, 32(10):1621-34.
- **Duan YY, Smith EP (2006).** Evaluating water quality using power priors to incorporate historical information.. *Environmetrics*;
- **Duan YY, Smith EP , Ye KY (2006).** Using power priors to improve the binomial test of water quality. *Journal of Agricultural, Biological and Environmental Statistics*

- FDA (2010) Guidance for the use of Bayesian statistics in medical device clinical trials.
- FDA (2010) Non-inferiority clinical trials: guidance for industry (draft).
- FDA (2013) Expedited Programs for Serious Conditions - Drugs and Biologics (draft).
- Fuquene JA, Cook D, Pericchi LR (2009). A case for robust Bayesian priors with applications to clinical trials. *Bayesian Analysis* 4, 817-846.
- **Gelman A (2006).** Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1, 515-533.
- Gsponer T, Gerber F, Bornkamp B, Ohlssen D, Vandemeulebroecke M, Schmidli H (2014) A practical guide to Bayesian group sequential designs. *Pharmaceutical Statistics* 13:71-80.
- Gsteiger S, Neuenschwander B, Mercier F, Schmidli H (2013) Using historical control information for the design and analysis of clinical trials with over-dispersed count data. *Statistics in Medicine*, 32(21):3609-22.
- Hasselblad V and Kong DF (2001). Statistical methods for comparison to placebo in active-control trials. *Drug Information Journal* 35: 435–449.
- Higgins JP, Whitehead A (1996) Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine* 15, 2733-2749
- Higgins JPT, Green S (eds) (2011). Cochrane handbook for systematic reviews of interventions, version 5.1.0. March 2011, www.cochrane-handbook.org
- **Higgins, Thompson, Spiegelhalter (2009.** A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society A*

- **Hobbs BP, Carlin BP, Mandrekar SJ, Sargent DJ (2011).** Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. *Biometrics* 67, 1047-1056.
- **Hobbs BP, Carlin BP, Sargent DJ (2013).** Adaptive adjustment of the randomization ratio using historical control data. *Clinical Trials* 10, 430-440.
- van Houwelingen HC, Arends LR and Stijnen T (2002) Advanced methods in meta-analysis: multivariate approach and meta-regression. Statistics in Medicine 21(4):589–624.
- Hueber W et al. (2012) Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* 61, 1693-1700.
- **Ibrahim JG, Chen MH (2000).** Power prior distributions for regression models. *Statistical Science* 15, 46-60.
- Lumley T (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 21(16): 2313–2324.
- Lu G, Ades AE (2004) Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 23(20): 3105–3124.
- Lu G, Ades AE (2006) Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 101(474): 447–459.

- **Morita S, Thall PF, Müller P (2008).** Determining the effective sample size of a parametric prior. *Biometrics* 64, 595-602.
- **Morita S, Thall PF, Müller P (2012).** Prior Effective Sample Size in Conditionally Independent Hierarchical Models. Bayesian Analysis 7(3). doi: 10.1214/12-BA720.
- **Neuenschwander B, Branson M, and Spiegelhalter DJ (2009).** A note on the power prior. Statistics in Medicine 28, 3562-3566.
- **Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter DJ (2010).** Summarizing historical information on controls in clinical trials. *Clinical Trials* 7, 5-18.
- Neuenschwander B (2011). From historical data to priors. ASA Proceedings
- Nixon RM, Bansback N, Brennan A (2007) Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Statistics in Medicine* 26(6):1237-54
- **O'Hagan A (1979).** On outlier rejection phenomena in Bayes inference. Journal of the Royal Statistical Society, Series B
- O'Hagan A, Forster J (2004) *Bayesian Inference, Kendall's Advanced Theory of Statistics Vol2B*. Chichester: Wiley.
- **O'Hagan A, Pericchi L (2012).** Bayesian heavy-tailed models and conflict resolution: a review. *Brazilian Journal of Probability and Statistics* 26, 372-401
- **Pocock SJ (1976).** The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases* 29, 175-188.

- Ryan L (1993) Using historical controls in the analysis of developmental toxicity data. *Biometrics*, 1126-1135.
- Salanti G, Higgins JP, Ades AE, Ioannidis JP (2008) Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* 17(3): 279–301.
- Salanti G, Dias S, Welton NJ, Ades AE, Golfinopoulos V, Kyrgiou M, Mauri D, Ioannidis JP (2010) Evaluating novel agent effects in multiple-treatments meta-regression. Statistics in Medicine 2010 29(23):2369-83.
- Schmidli H, Wandel S, Neuenschwander B (2013) The network meta-analytic-predictive approach to non-inferiority trials. *Statistical Methods in Medical Research* ;22(2):219-40.
- **Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B (2014).** Robust meta-analytic-predictive priors in clinical trials with historical control information. *Submitted*.
- **Spiegelhalter DJ, Abrams KR, Myles JP (2004).** *Bayesian Approaches to Clinical trials and Health-Care Evaluation*. Chichester: John Wiley and Sons.
- Viele K, Berry S, Neuenschwander B et al. (2013). Use of historical control data for assessing treatment effects in clinical trials. Pharmaceutical Statistics
- Witte S, Schmidli H, O'Hagan A, Racine A (2011) Designing a non-inferiority study in kidney transplantation: a case study. *Pharmaceutical Statistics*, 10, 427-432