RBesT / brms (Stan)

Bayesian evidence synthesis for clinical trials with RBesT

Sebastian Weber

Recent Advances in Meta-Analysis, Göttingen

2023-08-24

NOVARTIS

Acknowledgements

- Beat Neuenschwander
- Heinz Schmidli
- Anthony O'Hagen
- Satrajit Roychoudhury
- Simon Wandel
- Tim Friede

Outline

Introduction

Use of bayesian evidence synthesis in biostatistics
Applications in drug development

- **historical control** data
  - **sample size reduction in control group** while maintaining statistical power
  - aid in trial design to define true effect
  - aid in assessment of design parameters like variability
  - probability of success
- pediatric extrapolation
  - predicting pediatric outcomes based on adult data
    Are children like small adults?
  - combine discounted adult evidence with pediatric data
- historical treatment effect data (network meta-analysis)
  - support futility decisions at interim analysis
  - derivation of non-inferiority margins
  - sample size reduction for head-to-head comparison trials

Use of historical control data in clinical trials

Goal: reduce control group sample size while maintaining power

Design a (**future**) trial using synthesized evidence on control:

- Collect historical data from relevant literature *systematically*
- Evaluate heterogeneity of historical data
  - data quality
  - patient population
  - trial design
- Pre-specify trial protocol
  - what is the evidence used precisely?
  - how is the main analysis conducted?
- Document properties of trial design using historical evidence
  - type I error
  - power
- What to do about the unexpected case when the new trial and historical data do not agree?

`RBesT`: R Bayesian evidence synthesis tools

RBesT supports in *using* historical data for clinical trials

Implementing the Meta-Analytic-**Predictive** (MAP) approach:

1. Assess historical data for relevance
   - **exchangeability assumption justifiable?**
   - between-trial heterogeniety τ?
2. Run MAP analysis to obtain
   informative MAP prior in *parametric form*
   - Analyse historical data using MCMC: `gMAP`
   - Approximate MCMC MAP prior with parametric density: `automixfit`
   - Consider robustification: `robustify`
3. Evaluate frequentist design properties: `oc1S` or `oc2S`
4. Run final analysis: `postmix`

Ankylosing Spondylitis, *The Lancet*, 2013, (382)

Double-blinded PoC study to test *secukinumab* against *placebo*

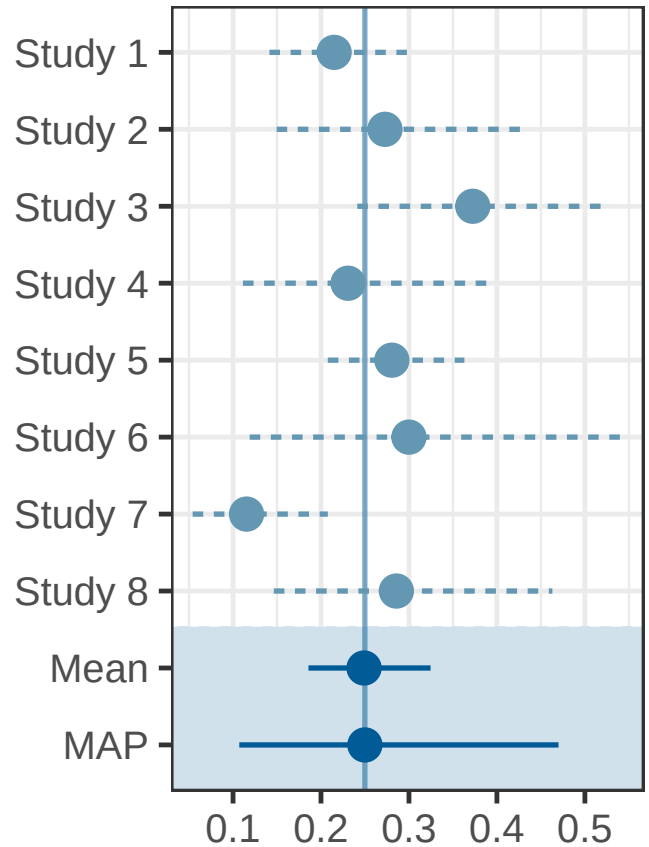
- **Endpoint**
  binary response ASAS20 at week 6
  higher rate is better
- **Bayesian design**
  - $Pr(\pi_t - \pi_p \geq 0 | y) > 0.95$
  - Placebo: MAP derived,

$$\pi_p \sim \text{Beta}(11, 32)$$

  - Active:

$$\pi_t \sim \text{Beta}(0.5, 1)$$

- **Randomization ratio**
  4:1 (24 vs 6)



NOVARTIS

Ankylosing Spondylitis, *The Lancet*, 2013, (382)
Data

The Ankylosing Spondylitiy data set is part of `RBesT` available upon loading the package as `data.frame` named `AS`:

study	n	r
Study 1	107	23
Study 2	44	12
Study 3	51	19
Study 4	39	9
Study 5	139	39
Study 6	20	6
Study 7	78	9
Study 8	35	10

```
  n   r
513 127
```

Prior derivation with gMAP

Generalized Meta-Analytic-Predictive model

Hierarchical model to obtain predictive of mean parameter

Y is the (control) group summary data for H historical trials

$$Y_h|\theta_h \sim f(\theta_h) \qquad \forall\ h \in [1, H]$$
$$Y_*|\theta_* \sim f(\theta_*) \qquad \text{for new trial (generative)}$$

Exchangeability assumption:

$$g(\theta_h)|\beta, \tau \sim \text{Normal}(\beta, \tau^2) \qquad \forall\ h \in [1, H]$$
$$g(\theta_*)|\beta, \tau \sim \text{Normal}(\beta, \tau^2) \qquad \text{for new trial (generative)}$$

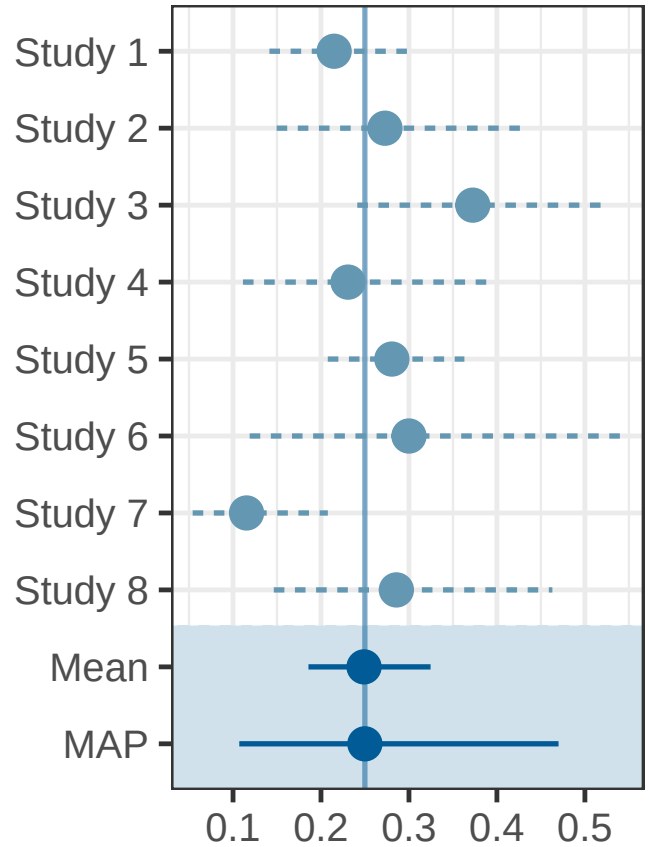
- f likelihood / g link function
  Binomial/logit, Normal (known σ)/identity or Poisson/log
- β population mean with prior Normal(m_β, s_β^2)
- τ between-trial heterogeneity with prior P_τ

Meta-Analytic-Predictive approach in words

A MAP prior is the predictive for the mean of a future trial

- **Mean**: $p(\beta|y)$ is the population mean or the *typical trial result*
- **MAP**: $p(\theta_*|y)$ is the *predictive distribution* for the mean of a *future trial*
  - Hierarchical model is *generative* as it extends to unseen trials
  - Equivalent: Future trial has no observed data yet, i.e. it's (to be collected) *data is missing*
- **Between-trial heterogeneity** τ critically governs borrowing:
  $\tau \to 0 \Rightarrow$ pooling
  $\tau \to \infty \Rightarrow$ stratification

Ankylosing Spondilytis example



NOVARTIS

The hierarchical model: A data driven prior

The normal-normal hierarchical model with known σ and τ with n_h measurements per group is:

$$y_h|\theta_h, \sigma \sim \text{Normal}(\theta_h, \sigma^2)$$
$$\theta_h|\beta, \tau \sim \text{Normal}(\beta, \tau^2)$$

Then the *conditional* posterior on y_h for θ_h is (β & τ known):

$$\theta_h|\beta, \tau, y_h \sim \text{Normal}(\hat{\theta}_h, V_h)$$

$$\hat{\theta}_h = \frac{\frac{1}{\tau^2}\beta + \frac{1}{se_h^2}\bar{y}_h}{\frac{1}{V_h}} \text{ and } \frac{1}{V_h} = \frac{1}{\tau^2} + \frac{1}{se_h^2}$$

The per-group mean $\hat{\theta}_h$ is a precision weighted average of the data-mean $\bar{y}_h$ and the population mean β

Conservative prior choices for τ and β

Binary and normal endpoints

Endpoint		very conservative[1] τ prior	conservative[1,2] τ prior	β prior[3]
Binary	$0.2 < \pi < 0.8$	$\mathrm{N}^+(0, 1)$	$\mathrm{N}^+(0, (1/2)^2)$	$\mathrm{N}(0, 2^2)$
Normal	known σ	$\mathrm{N}^+(0, (\sigma/2)^2)$	$\mathrm{N}^+(0, (\sigma/4)^2)$	$\mathrm{N}(\mu_0, \sigma^2)$

1. very conservative, see *Neuenschwander et al., 2010*
2. less heterogeneous data as often seen empirically in meta-analysis, see *Friede et al., 2016*
3. unit-information prior for β (single observation of no effect), see *Kass & Wasserman, 1995*
   μ_0 set problem dependent (often 0)

Refer to *Röver et al., 2021* for an in-depth discussion on priors for τ

Running the MAP analysis with gMAP
Ankylosing Spondylitis example

```
set.seed(123234)
map_mc <- gMAP(cbind(r, n-r) ~ 1 | study, data=AS, family=binomial,
               tau.dist="HalfNormal", tau.prior=1, beta.prior=2)
```

- `set.seed` ensures exact reproducibility
- model formula follows standard R conventions
  `cbind(#responders, #non-responders) ~ 1 | study`
- data-set `AS` (part of `RBesT`) passed in as `data.frame`
- `family` selects likelihood (and link function)
- τ prior *must* be set (very conservative here)
- β prior *should* be given (very conservative here)

gMAP results for Ankylosing Spondylitis analysis

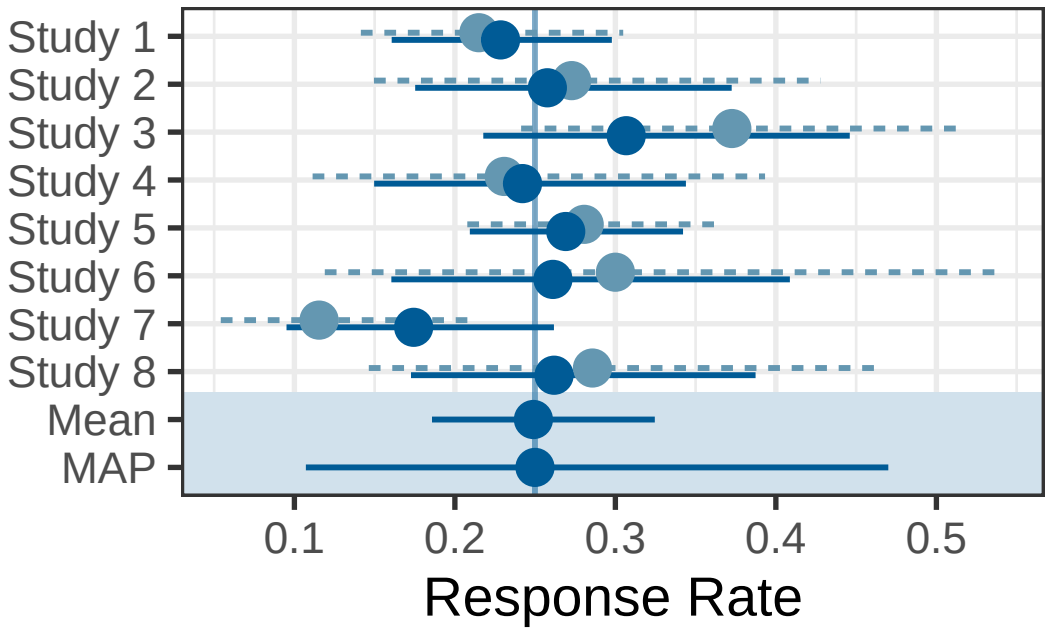
```
print(map_mc)
```

```
Generalized Meta Analytic Predictive Prior Analysis

Call:  gMAP(formula = cbind(r, n - r) ~ 1 | study, family = binomial,
    data = AS, tau.dist = "HalfNormal", tau.prior = 1, beta.prior = 2)

Exchangeability tau strata: 1
Prediction tau stratum    : 1
Maximal Rhat              : 1

Between-trial heterogeneity of tau prediction stratum
  mean     sd   2.5%    50%  97.5%
0.3750 0.2130 0.0407 0.3500 0.8720

MAP Prior MCMC sample
  mean     sd   2.5%    50%  97.5%
0.2580 0.0875 0.1070 0.2500 0.4700
```

Graphical model diagnostics

Standard forest plot with meta-analytic model estimates

```
# Ankylosing Spondylitis example
plot(map_mc, size=0.5)$forest_model
```



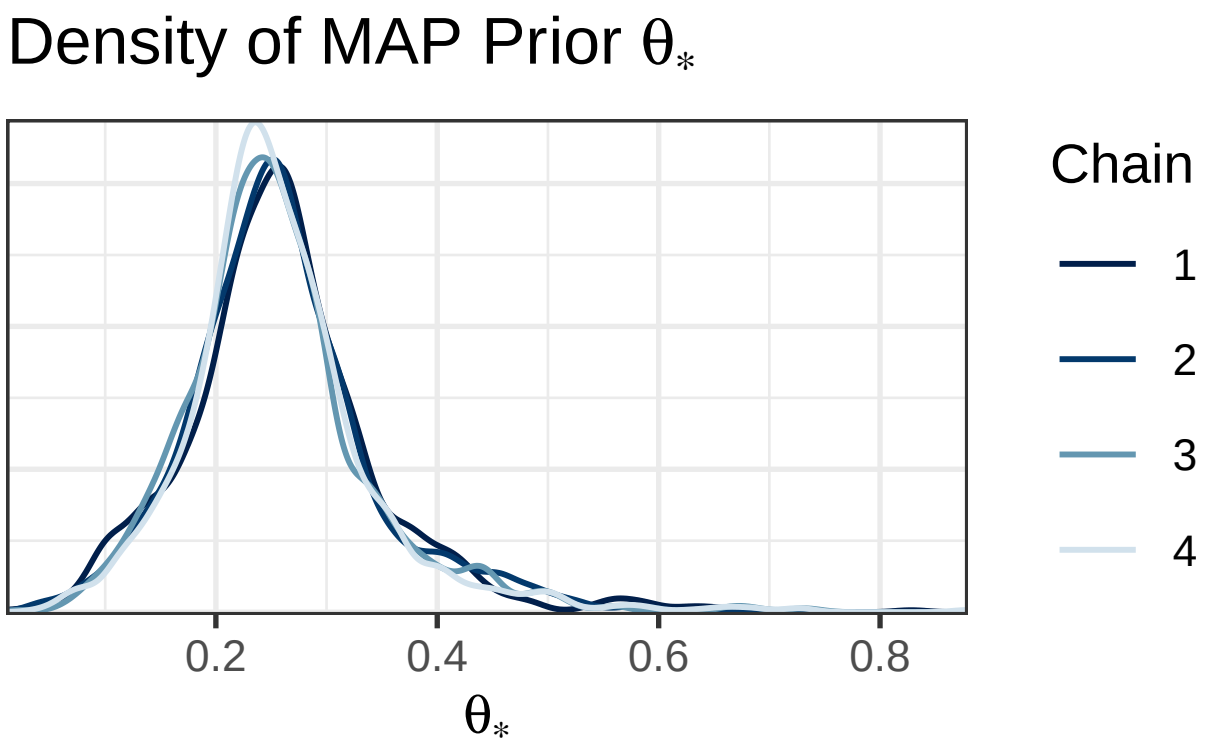
- `forest_plot` function produces customizable forest plots
- please refer to "Customizing RBesT Plots" vignette for basic customization of `ggplot2` plots

NOVARTIS

Using MAP priors for clinical trials

MAP analysis result is a MCMC sample

```
plot(map_mc)$densityThetaStar
```



A MCMC sample of 4×10^3 draws is unconvenient to communicate or pre-specify in a protocol

Turning MAP into a parametric density

Parametric densities have many practical advantages

- **Conjugate** priors allow for **fast analytic** manipulations
  the posterior is then given by the same distributional class as the prior

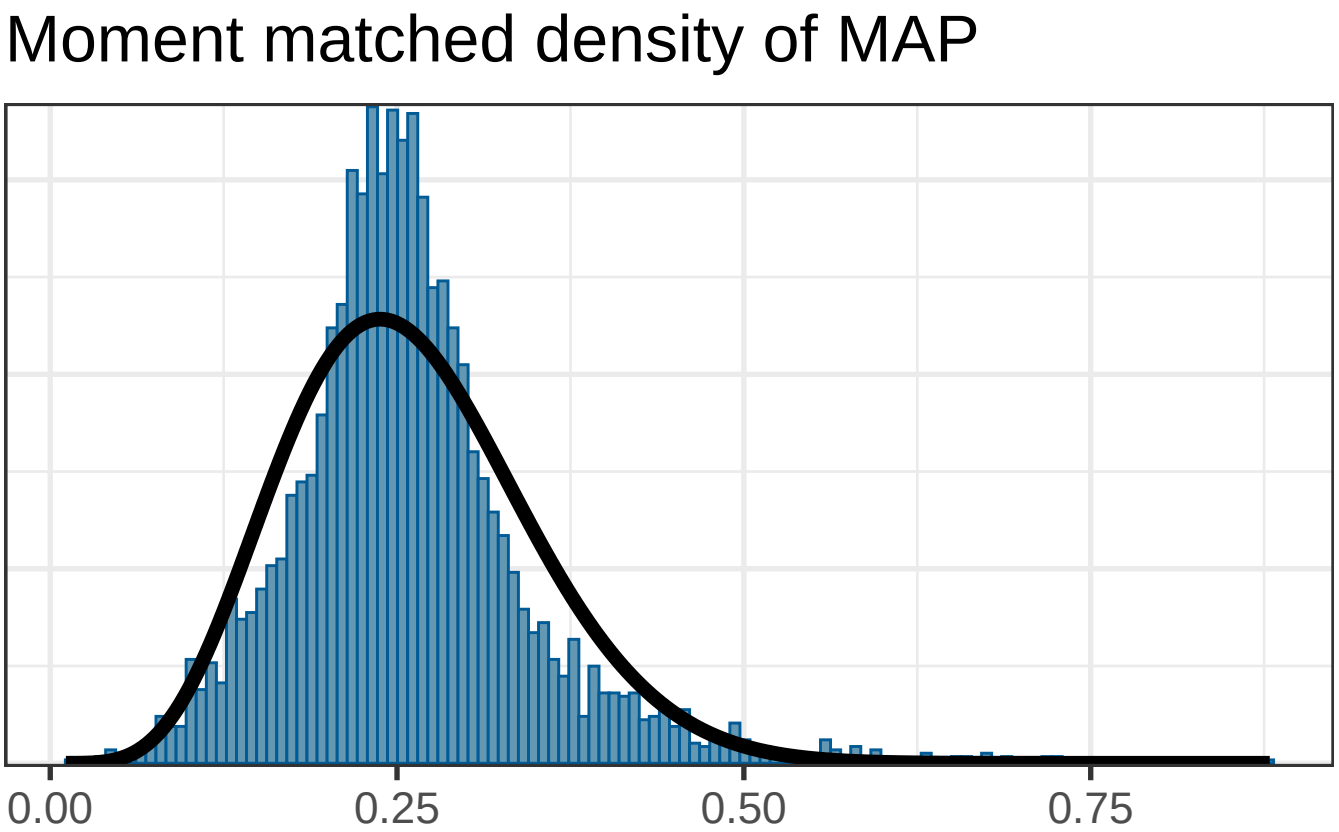
Likelihood	Prior	Posterior
Binomial	Beta	Beta
Normal (known σ)	Normal	Normal
Poisson	Gamma	Gamma

- Simple moment matching often not accurate (heavy tails)
  $\Rightarrow$ mixtures are arbitrarily accurate and maintain conjugacy

Simple moment matching inaccurate

Heavy tails of MAP priors lead to misfit

```
map_moment_match <- mixfit(map_mc, Nc = 1)
plot(map_moment_match)$mix + ggtitle("Moment matched density of MAP")
```

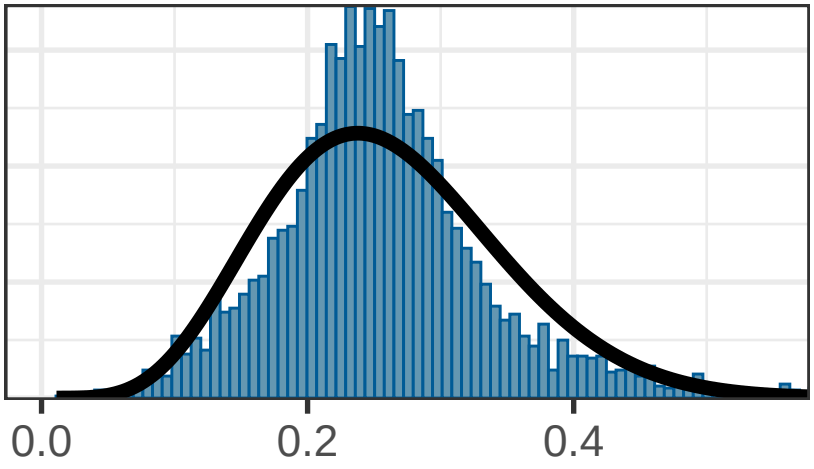


NOVARTIS

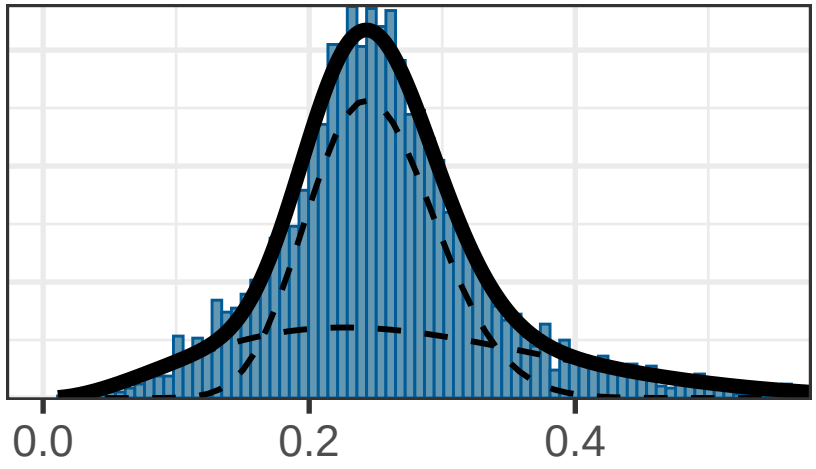
Mixtures improve accuracy of parametric MAP

Mixture MAP prior is an accurate parametric approximation

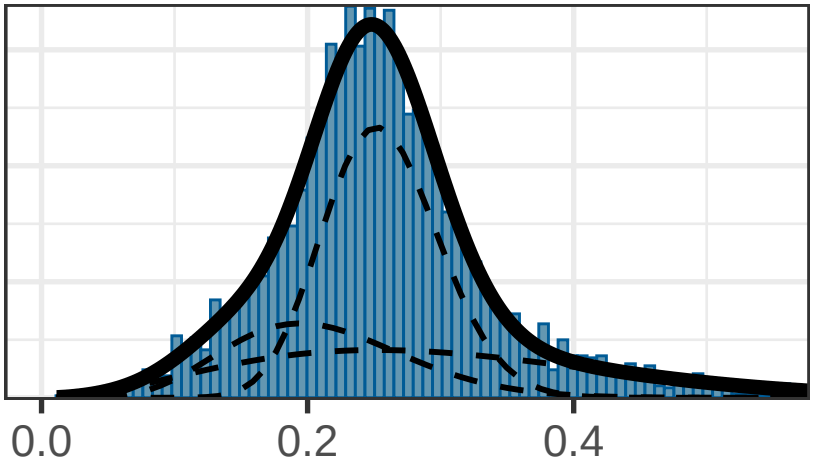
Beta prior (moment matched)



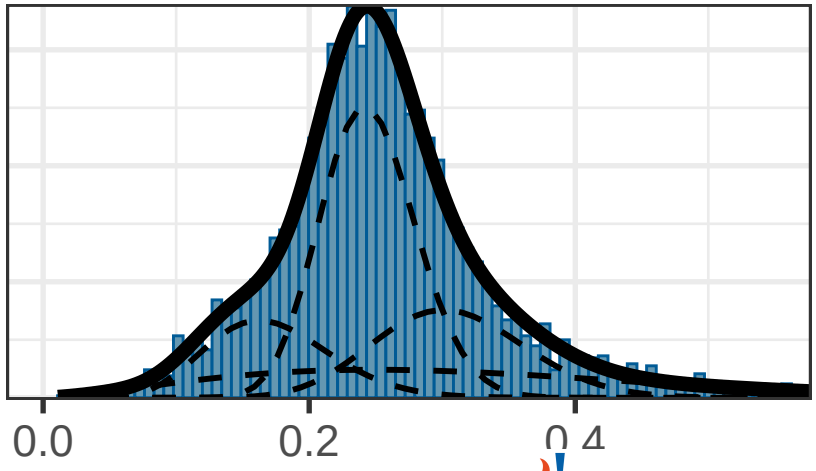
2−comp. Beta mixture



3−comp. Beta mixture



4−comp. Beta mixture



NOVARTIS

Mixtures improve accuracy of parametric MAP
Automatic AIC based selection for number of components

```
# EM used to fit 1-4 components and selects model via AIC
map_automix <- automixfit(map_mc)
print(map_automix)
## EM for Beta Mixture Model
## Log-Likelihood = 4462
##
## Univariate beta mixture
## Mixture Components:
##    comp1  comp2  comp3
## w  0.506  0.281  0.213
## a 25.962  3.245  7.744
## b 75.145  7.668 28.916
```

```
# Recommended graphical check of EM fit:
plot(map_automix)$mix
```

Design planning, operating characteristics

Evaluating trial designs classically
Binary responder analysis

- Type I error α for no effect hypothesis H_0
  $\theta_p = \theta_t$
- Sample size per group N_p & N_t chosen under true effect assumption (alternative, H_a) and desired type II error β
  θ_p, $\theta_t = \theta_p + \delta$
- 1:1 randomization has highest efficiency for two-arm trial
  $N_p = N_t$
- Type I error to reject H_0 is controlled at α
  - for any θ
  - for any sample size N

Evaluating trial designs with RBesT
Binary responder analysis (Ankylosing Spondylitis example)

```
alpha <- 0.05
## 1. Define decision criterium for success
##    here: 2-sample decision criterium, P(p_treat - p_placebo > 0) > 0.95
decision <- decision2S(1 - alpha, 0, lower.tail=FALSE)

## 2. Define design (priors, sample size, decision)
uniform_prior <- mixbeta(c(1, 1, 1))
design_uniform_classic <- oc2S(uniform_prior, uniform_prior, 24, 24, decision)

## 3. Evaluate power
design_uniform_classic(0.25+0.35, 0.25)
```

```
[1] 0.819
```

```
## For comparison the respective frequentist calculation
power.prop.test(24, 0.25+0.35, 0.25, sig.level=0.05, alt="one.sided")$power
```

```
[1] 0.806
```

Evaluating trial designs with RBesT using MAP
Using MAP priors allows to reduce (control) sample size

- Informative MAP priors enable unequal randomization by *substituting* sample size of the control by prior information
- An informative prior can be considerd to have an *effective sample size* (ess)
- Robustification of MAP prior strongly recommended
  $\Rightarrow$ reduces ess of MAP prior

```
## Ankylosing Spondylitis example
## 0. derive MAP prior
## 1. use classic operating characeristics to determine starting sample size
## 2. reduce control group sample size by about ess samples

map_robust  <- robustify(map_automix, weight=0.2, mean=0.5)
round(ess(map_robust))
```

```
[1] 27
```

So we may substantially reduce the control group here!

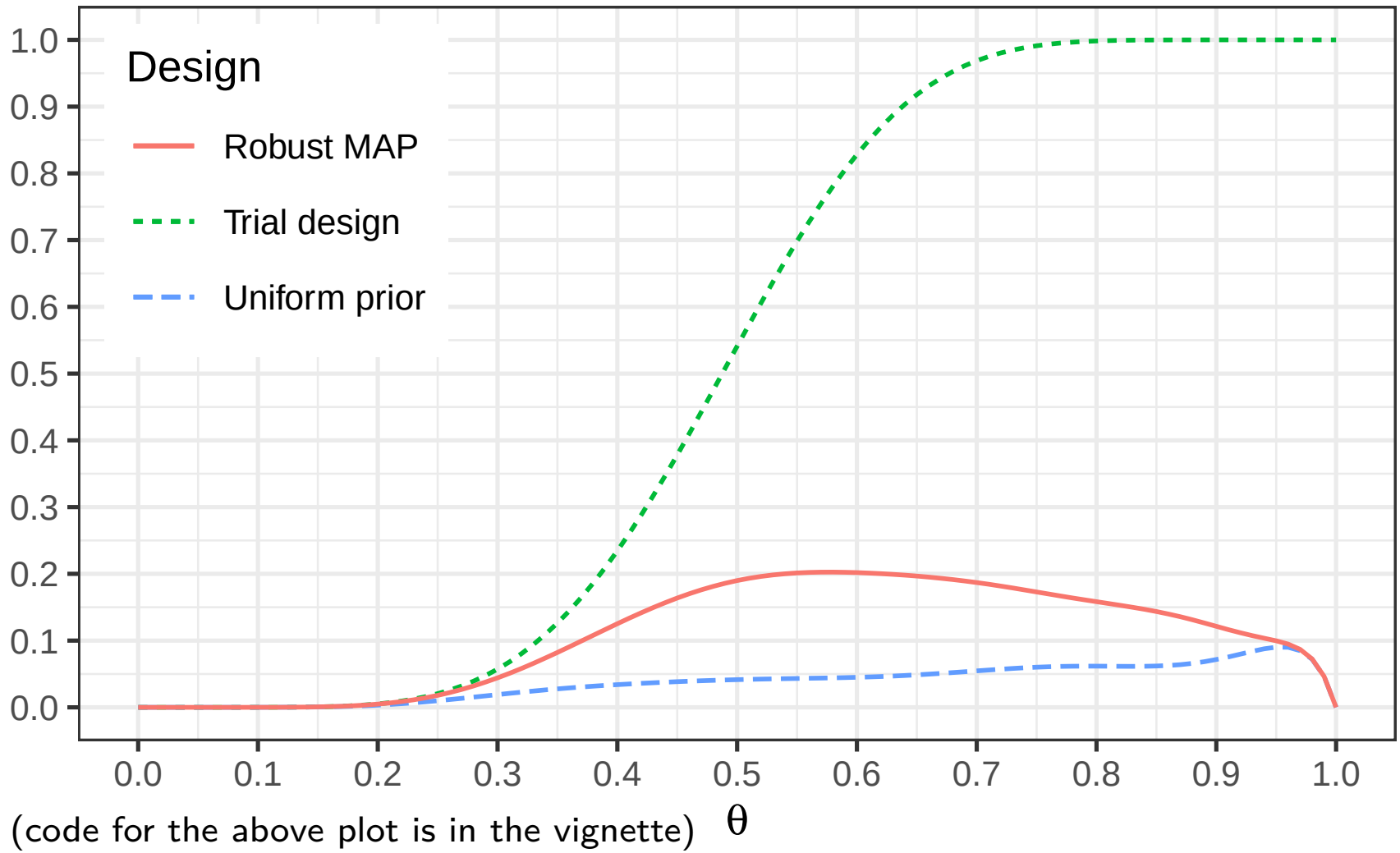
Operating Characteristics
Ankylosing Spondylitis example

For $N_p = 6$ and $N_t = 24$
First definition of design, then exact calculations (binary case)

```
# Define decision criterium, P(p_treat - p_placebo > 0) > 0.95
decision <- decision2S(0.95, 0, lower.tail=FALSE)

treat_prior   <- mixbeta(c(1, 0.5, 1)) # Prior for treatment arm
placebo_prior <- mixbeta(c(1,11  ,32)) # Prior for placebo arm as used
uniform_prior <- mixbeta(c(1, 1  , 1)) # Uniform prior for comparison
rmap_prior    <- robustify(map_automix, weight=0.2, mean=0.5) # robust MAP

# Calculate design properties (depends on priors, sample size & decision)
design_uniform <- oc2S(uniform_prior, uniform_prior, 24,  6, decision)
design_trial   <- oc2S(treat_prior,   placebo_prior, 24,  6, decision)
design_robust  <- oc2S(treat_prior,   rmap_prior   , 24,  6, decision)

# Note: decision functions take mixtures as arguments and return
# 0="NO GO", 1="GO"
decision(postmix(treat_prior, r=15, n=24), postmix(rmap_prior, r=1, n=6))
## [1] 1
```

Type I Error, Frequency of GO for $\theta = \theta_t = \theta_p$

Ankylosing Spondylitis example

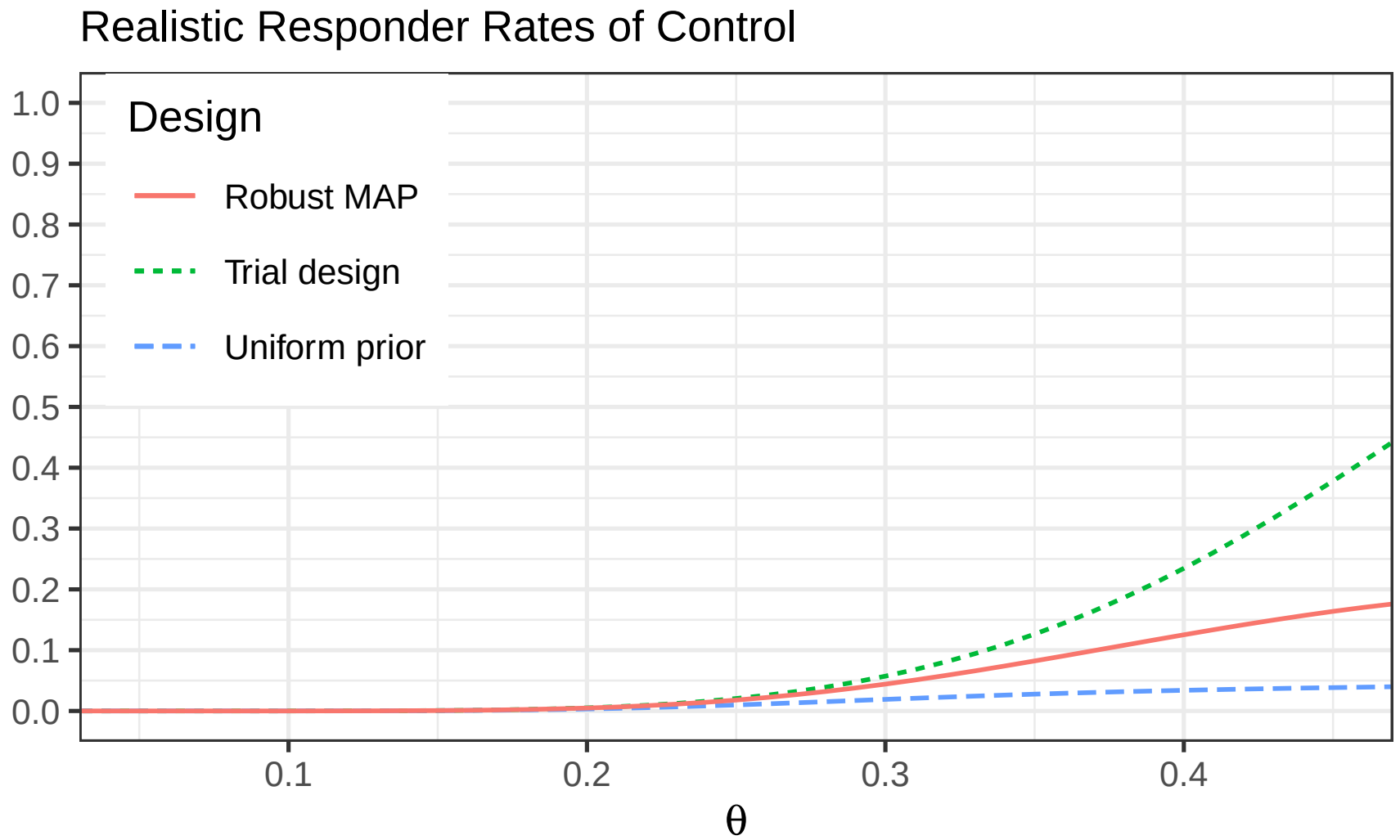


(code for the above plot is in the vignette)

NOVARTIS

Type I Error, Frequency of GO for $\theta = \theta_t = \theta_p$

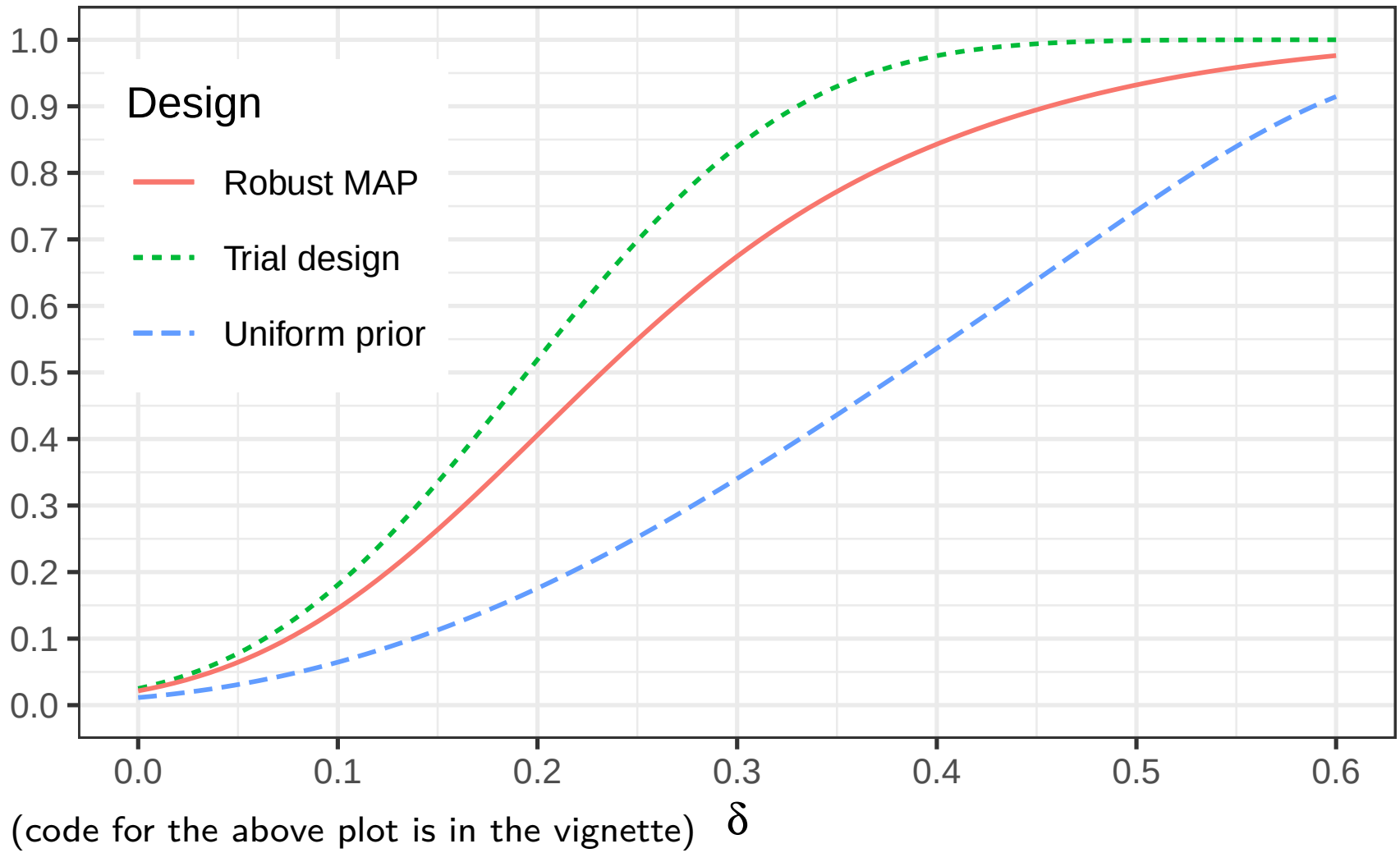
Ankylosing Spondylitis example



NOVARTIS

Power, Frequency of GO for $\theta_t = \theta_p + \delta$ $(\theta_p = \bar{\theta}_p)$

Ankylosing Spondylitis example

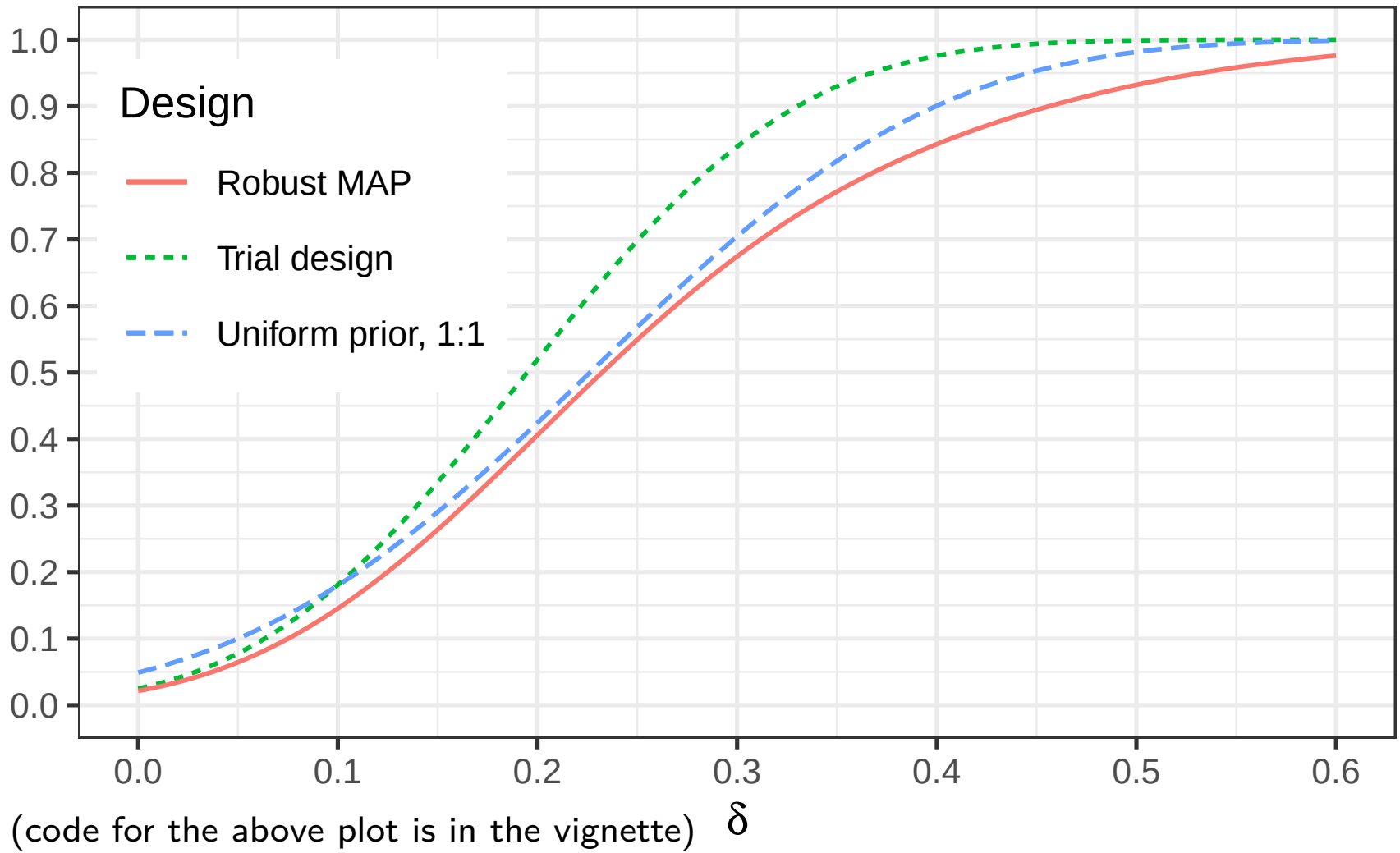


(code for the above plot is in the vignette)

NOVARTIS

Power, Frequency of GO for $\theta_t = \theta_p + \delta$ $(\theta_p = \bar{\theta}_p)$

Ankylosing Spondylitis example



(code for the above plot is in the vignette)

NOVARTIS

Useful resources

Useful resources

- Install RBesT
  `R> install.packages("RBesT", dependencies=TRUE)`
- Using RBesT
  `R> library(RBesT)`
- RBesT R help
  `R> ?gMAP` for help on gMAP
- **Vignettes** are easy to follow step-by-step introductions
- Homepage
  https://opensource.nibr.com/RBesT/
- CRAN
  https://CRAN.R-project.org/package=RBesT
- GitHub
  https://github.com/Novartis/RBesT

Getting RBesT and Help
RBesT is integrated into the R system

- Inter-linked HTML pages with `help.start()`
  opens a web-browser or RStudio help then follow
  `Packages -> RBesT`
- PDF reference distributed with RBesT (LaTeX formulas)
- **Vignettes**, see https://cran.r-project.org/package=RBesT
  - introduction: Getting started (binary endpoint)
  - introduction normal: Getting started (normal endpoint)
  - customizing plots: Plotting help
  - robustMAP: Reproduces Schmidli et al. (2014)
  - ...

Executive Summary

Executive Summary

RBesT: R Bayesian evidence synthesis tools

- Facilitates the application of the Meta-Analytic-Predictive (MAP) approach in clinical trials
- `RBesT` is designed as a modern R library documented, R integration, unit tested, fast & accurate
- Supports **binary, normal (known σ) and Poisson** endpoints (piecewise-constant time-to-event via Poisson)
- High abstraction level makes (complex) computations straightforward and user-friendly (trial statistician-friendly)

Weber, Li, Seaman, Kakizume & Schmidli (2021) JSS, 100, pp. 1–32. doi: 10.18637/jss.v100.i19

Backup

Effective sample size of a (MAP) prior

Measure of informativeness

- Conceptually for a Normal endpoint with known σ, the information of a simple (no mixture) Normal prior is

$$se = \frac{\sigma}{\sqrt{\text{ESS}}} \Leftrightarrow \text{ESS} = \frac{\sigma^2}{se^2}$$

- Available methods in `RBesT`
  - Moment based matching
  - Morita et al. (2008) based on curvature at mode of prior
  - **ELIR**, Neuenschwander et al. (2020)
- In-depth presentation: https://youtu.be/WgvrQZqGP9U

Effective sample size of a (MAP) prior
Measure of informativeness

ESS_{elir} is the default method in `RBesT` as it is the only ESS method which is **predictivley consistent**[†]**:**

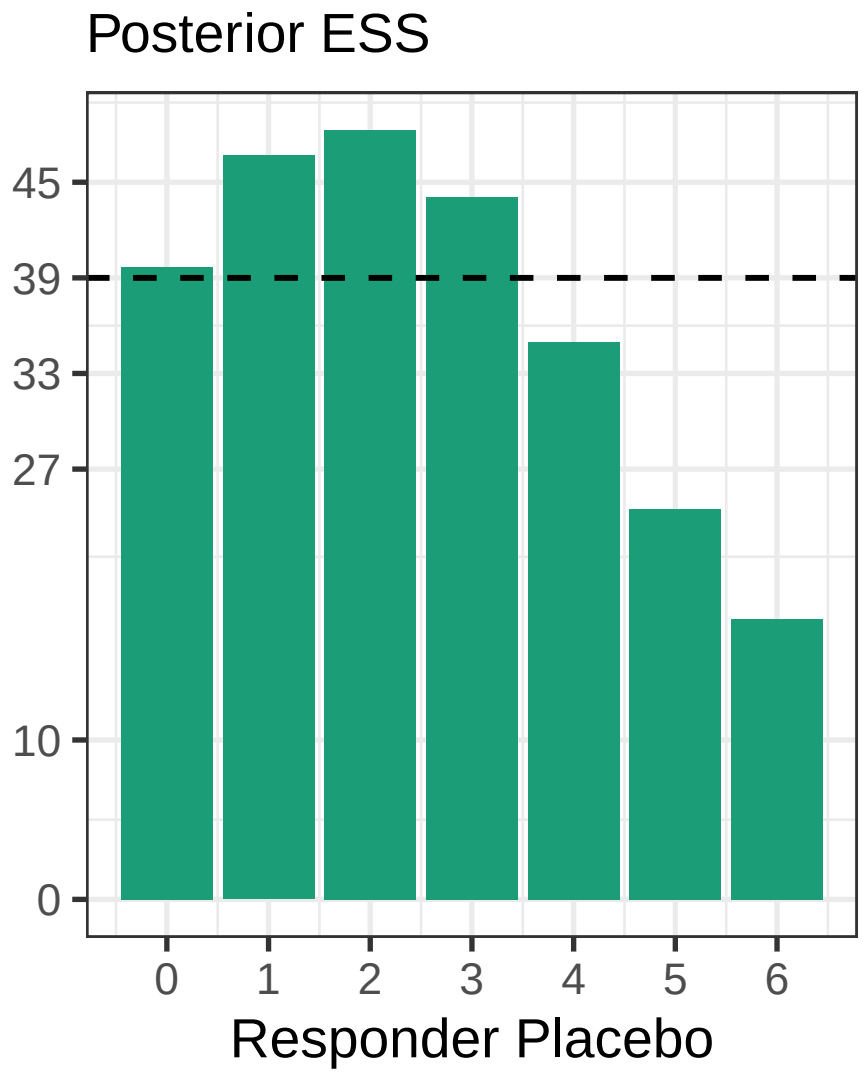
> *"For a sample size of M, the expected posterior ESS must be the sum of the prior ESS and M."*

$$E_{Y_M}\left[ESS\{p(\theta|Y_M)\}\right] = ESS\{p(\theta)\} + M$$

- ESS **applicable to any density - prior and posterior!**
- **Expected ESS of posterior must be consistent wrt to prior predictive distribution.**
- ELIR: "Expected local information ratio" (average curvature)

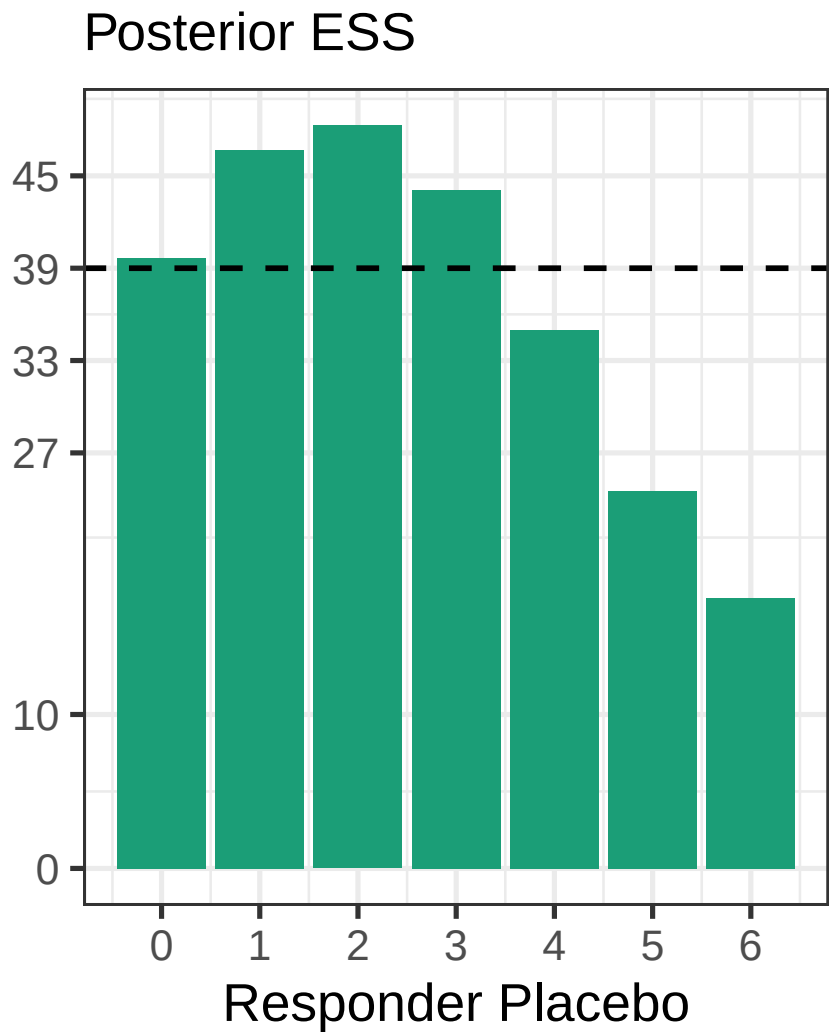
[†] Neuenschwander, Weber, Schmidli & O'Hagan (2020), Biometrics 76(2), pp. 578–587. doi: 10.1111/biom.13252.

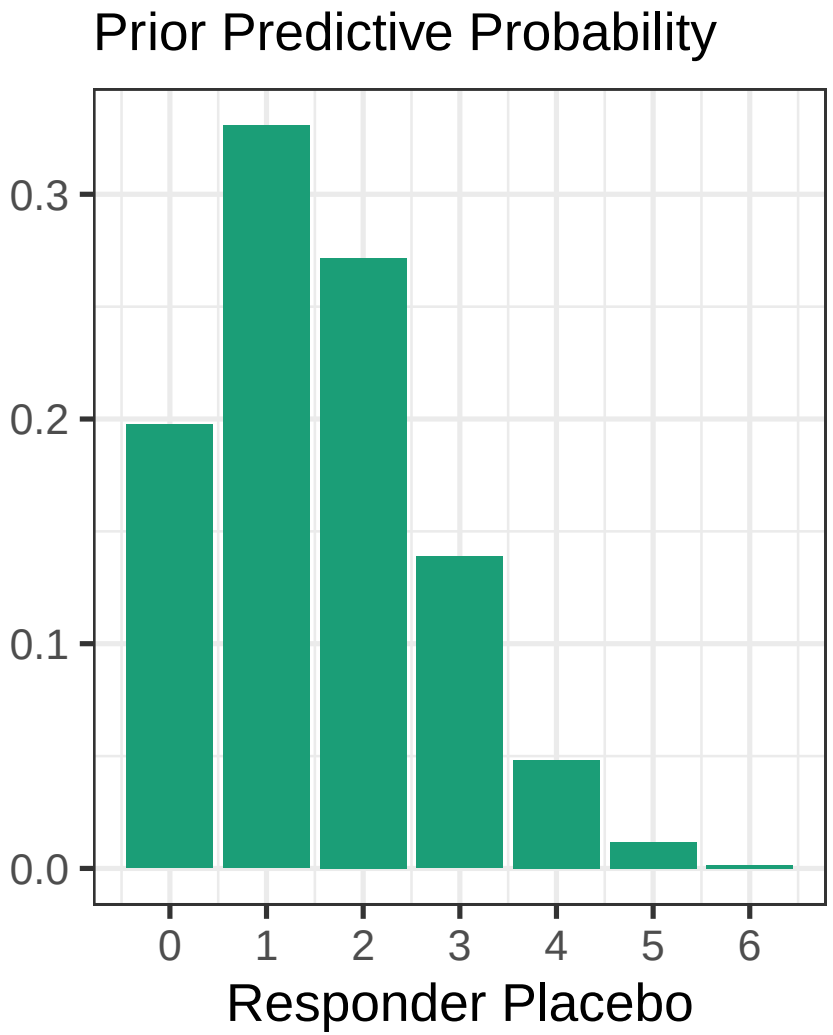
Ankylosing spondylitis trial *data scenarios*
MAP prior ESS 39



NOVARTIS

Ankylosing spondylitis trial *data scenarios*
MAP prior ESS 39

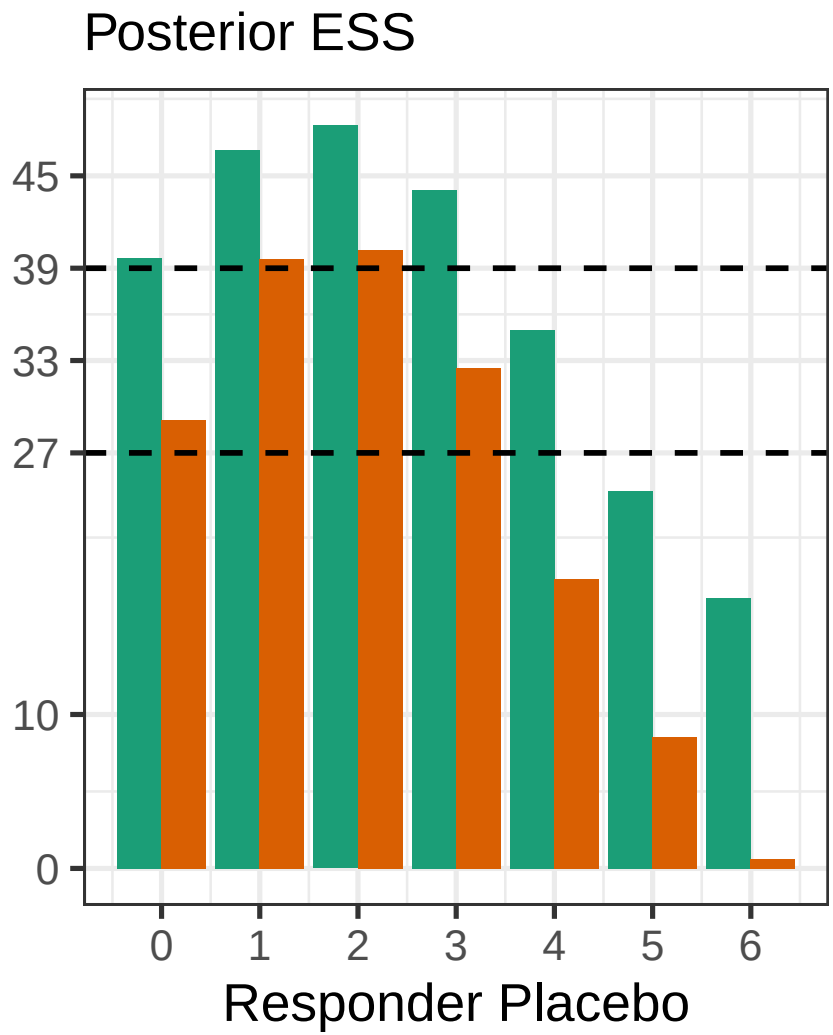


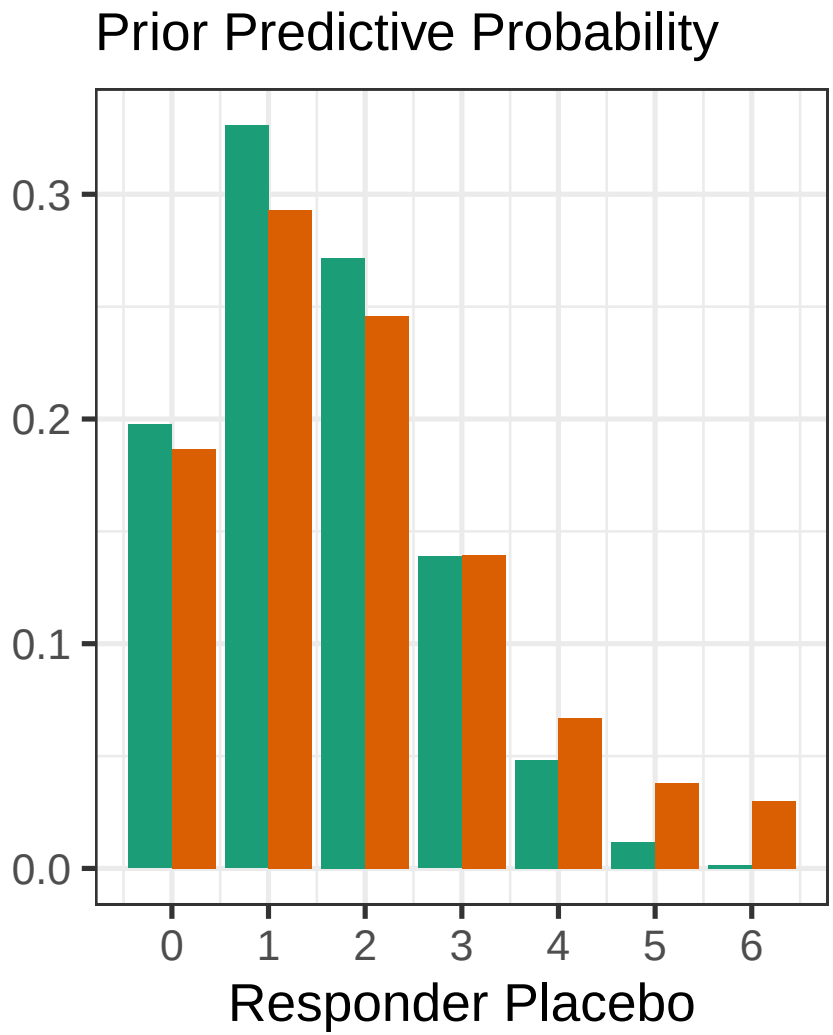


NOVARTIS

Ankylosing spondylitis trial *data scenarios*

MAP prior ESS 39 (green), robust MAP prior ESS 27 (red)





NOVARTIS

ESS predictive consistency robust MAP prior

r_p	n	PriorPred	Ess	Ess * PriorPred
0	6	0.19	29.14	5.44
1	6	0.29	39.60	11.60
2	6	0.25	40.16	9.87
3	6	0.14	32.54	4.54
4	6	0.07	18.83	1.26
5	6	0.04	8.55	0.32
6	6	0.03	0.59	0.02

```
sum(ess_rmap_posterior$Ess * ess_rmap_posterior$PriorPred)
```

```
[1] 33.1
```

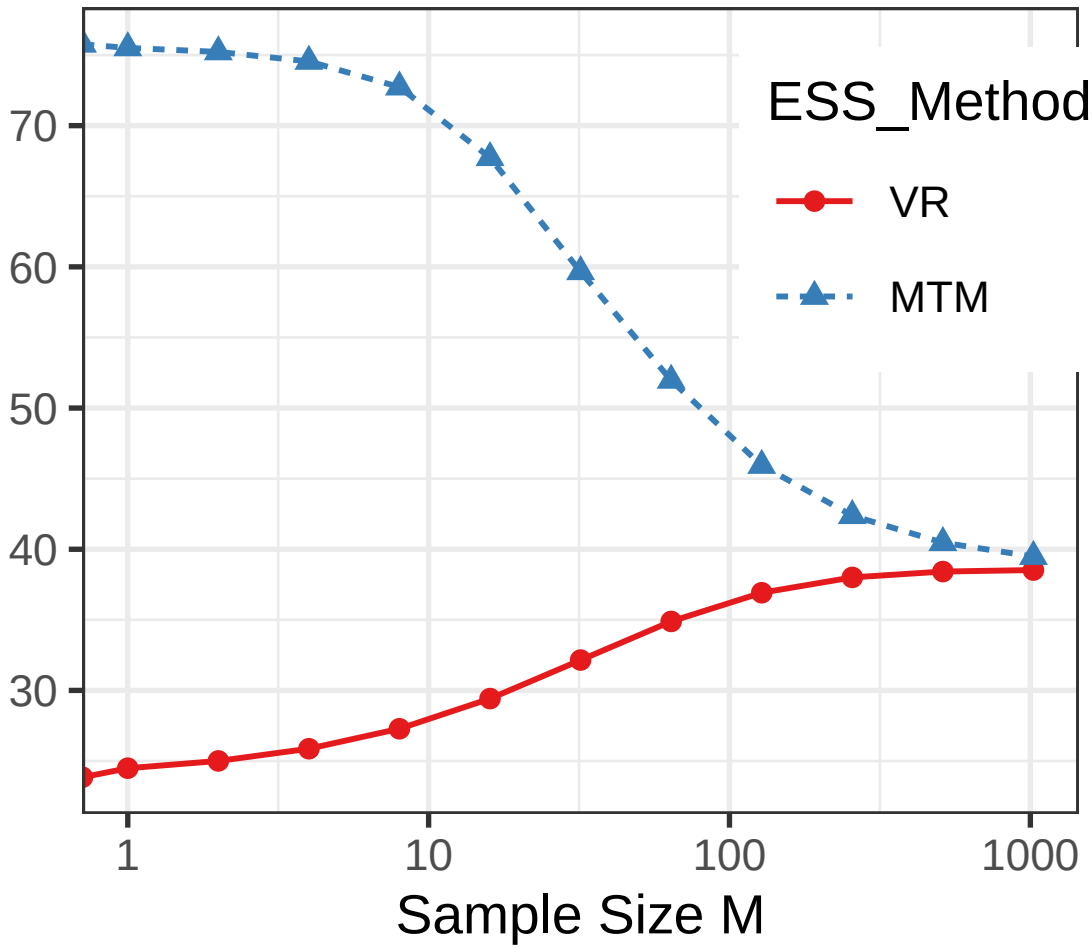
```
ess(rmap) + 6
```

```
[1] 33.1
```

$$E_{Y_6}\left[ESS\{p(\theta|Y_6)\}\right] = ESS\{p(\theta)\} + 6$$

Predictive (in-)consistency of ESS methods

$$E_{Y_M}\left\{ESS\left(p(\theta|Y_M)\right)\right\} - M$$

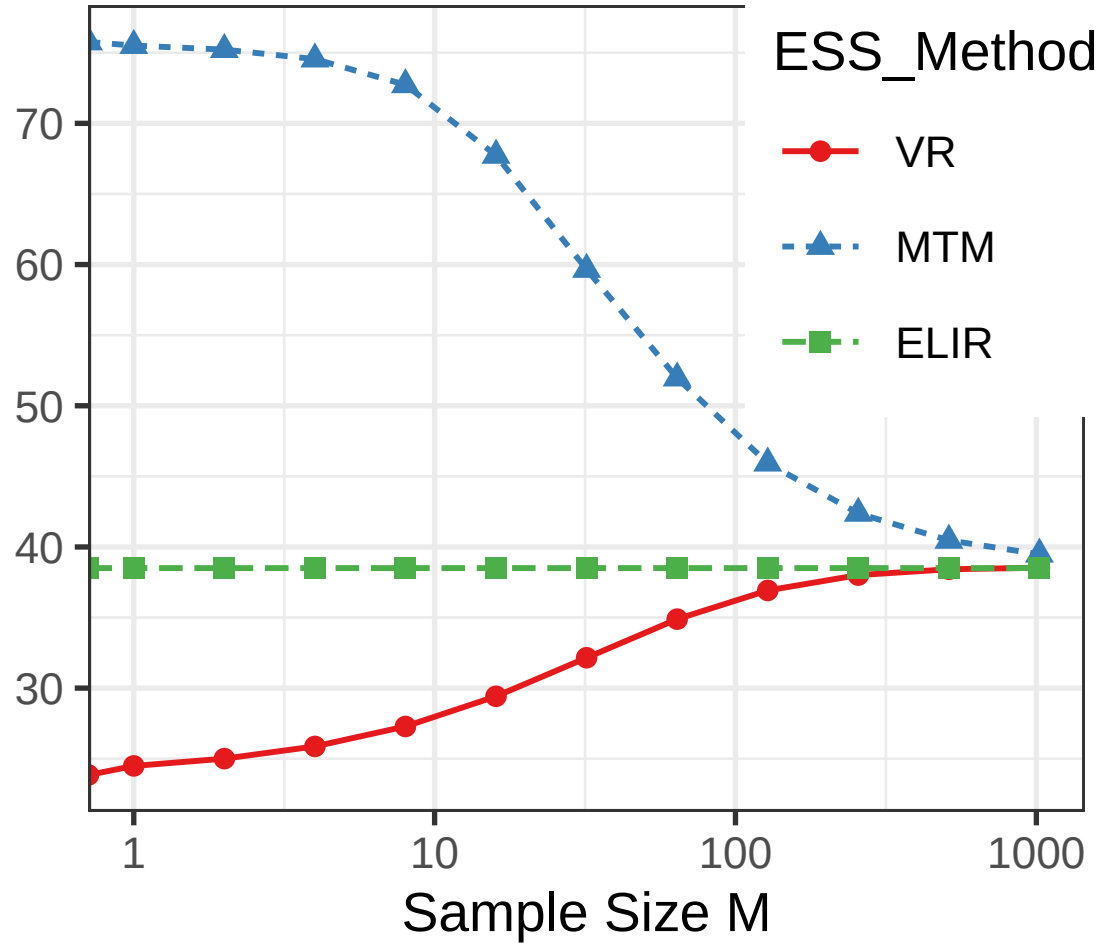


Example: MAP prior of Ankylosing Spondylitis data set.

NOVARTIS

Predictive (in-)consistency of ESS methods

$$E_{Y_M}\left\{ESS\left(p(\theta|Y_M)\right)\right\} - M$$



Predictive consistency holds "in expectation wrt to *prior predictive*".

NOVARTIS

MAP analyses are 2-step by construction

1. Retrospective analysis of historical data y_H to derive MAP prior for new trial
$$p(\theta_\star|y_H)$$
2. Analysis of new data $y_\star$ using MAP prior $p(\theta_\star|y_H)$ to obtain final posterior
$$p(\theta_\star|y_\star, y_H)$$

MAP analyses can be done in a single step

Sequential updating can be done in a single step as guranteed by Bayes' rule.
"Meta-Analytic-Combined" MAC analysis:

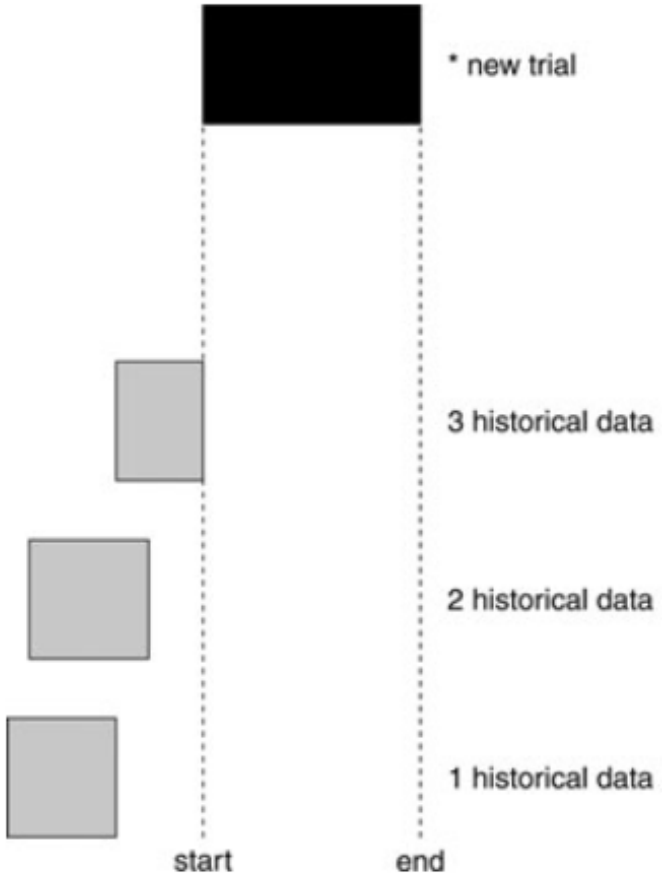
$$p(\theta_\star|y_\star, y_H) \propto p(\theta_\star|y_\star)\, p(\theta_\star|y_H)$$

This holds under the meta-analytic model (**no robustification**).

Proof is based on independence of y_H and $y_\star$.

See appendix of "PoS for Co-Data" vignette or Schmidli et al. (2014)

MAP analyses historical data retrospectively

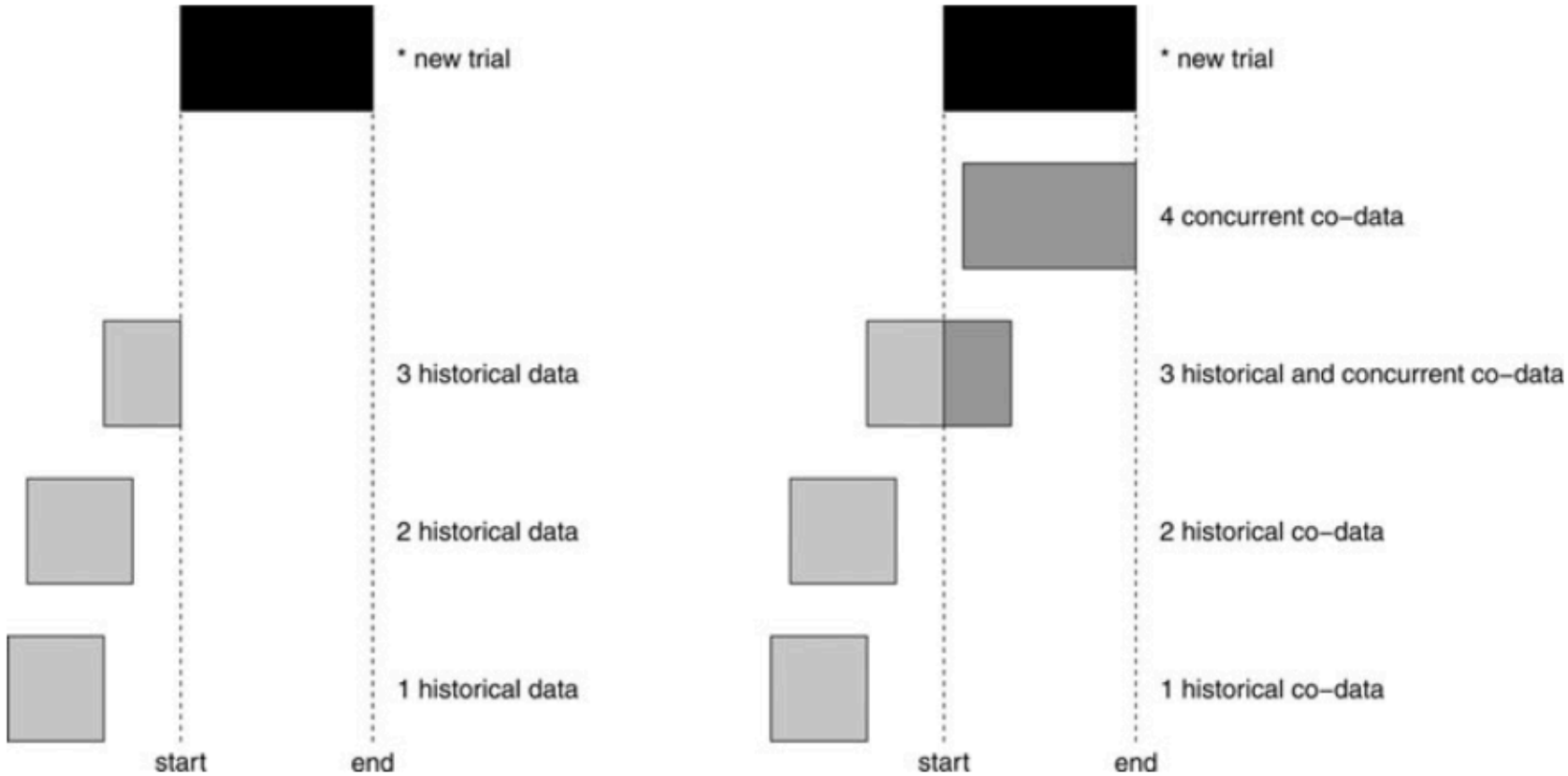


Leverage information from the past.

See Neuenschwander, et al. (2016), doi: 10.1080/19466315.2016.1174149.

NOVARTIS

MAC analyses historical & concurrent data jointly



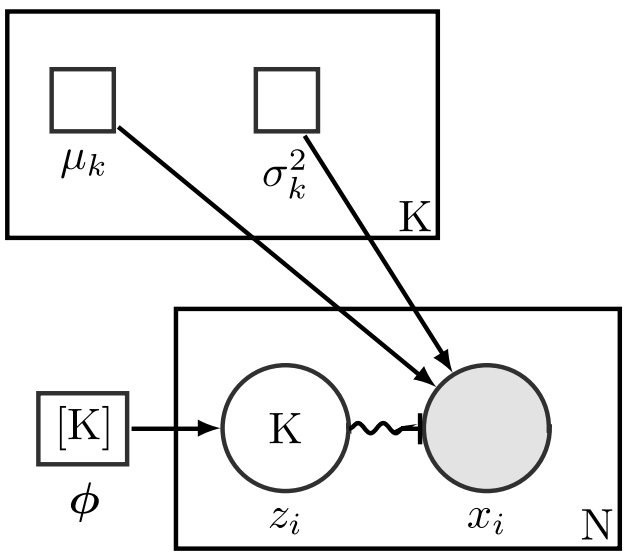
Leverage information from the past **and** present.

See Neuenschwander, et al. (2016), doi: 10.1080/19466315.2016.1174149.

NOVARTIS

Mixture Models
Estimation with Expectation-Maximization (EM)

$$\log p(x|w,a,b) = \sum_{n=1}^{N} \log \left[\sum_{k=1}^{K} w_k \, p(x_n|a_k,b_k) \right]$$



Example: Univariate normal

Source: Wikipedia

EM "trick" is to extend the likelihood

$$p(x|w,a,b) = \int p(x,z|w,a,b)\, dz$$

- x **observed** data as recorded
- z **latent** data, i.e. component indicator
- (x,z) **complete** data

NOVARTIS

Posterior Analysis for Mixture Priors

Fixed prior weights change in the posterior

Assume a mixture prior for some parameter θ

$$p(\theta, \mathbf{w}, \mathbf{a}, \mathbf{b}) = \sum_{k=1}^{K} w_k \, p_k(\theta, a_k, b_k)$$

for data y and likelihood $f(y|\theta)$, then the posterior is again a mixture equal to the posterior of each component and updated weights

$$p(\theta, \mathbf{w}, \mathbf{a}, \mathbf{b}|y) = \sum_{k=1}^{K} w'_k \, p_k(\theta, a_k, b_k|y)$$

Note: The prior weights w_k, are not random (fixed) but are still updated to $w'_k = w^*_k / \sum_{k=1}^{K} w^*_k$ with

(marginal likelihood) $w^*_k \equiv w_k \int f(y|\theta) \, p_k(\theta, a_k, b_k) \, d\theta = w_k \, p_k(y)$

Analytic Operating Characteristics in RBesT

RBesT calculates OCs for one-sided designs fast

The decision function $D(y_1, y_2)$, priors and sample sizes uniquely define the **decision boundary** $D_1(y_2)$ (conditional critical values):

$$D_1(y_2) = \sup_{y_1}\{D(y_1, y_2) = 1\},$$

$$\iint f_1(y_1|\theta_1)\, D(y_1, y_2)\, f_2(y_2|\theta_2)\, dy_1 dy_2 = \int F_1(D_1(y_2)|\theta_1)\, f_2(y_2|\theta_2)\, d$$

$D_1(y_2)$ is calculated when calling oc2S. Then all calls to the returned function evaluate the frequency for 1 assuming that y_1 (y_2) is distributed according to the assumed true value of θ_1 (θ_2).

Binary case calculation is exact, other endpoints use adaptive quadrature integration.

References I

Baeten, D., Baraliakos, X., Braun, J., Sieper, J., Emery, P., Heijde, D. van der, et al. (2013) Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: A randomised, double-blind, placebo-controlled trial. *The Lancet*, **382**, 1705–1713.

Dalal, S.R. and Hall, W.J. (1983) Approximating Priors by Mixtures of Natural Conjugate Priors. *Journal of the Royal Statistical Society. Series B (Methodological)*, **45**, 278–286.

Neuenschwander, B., Capkun-Niggli, G., Branson, M. and Spiegelhalter, D.J. (2010) Summarizing historical information on controls in. *Clin Trials*, **7**, 5–18.

Neuenschwander, B., Weber, S., Schmidli, H. and O'Hagan, A. (2020) Predictively consistent prior effective sample sizes. *Biometrics*, **76**, 578–587.

Röver, C., Bender, R., Dias, S., Schmid, C.H., Schmidli, H., Sturtz, S., et al. (2021) On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, **12**, 448–474.

Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D. and Neuenschwander, B. (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biom*.

References II

Spiegelhalter, D.J., Abrams, K.R. and Myles, J.P. (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation: Spiegelhalter/Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Ltd, Chichester, UK.

Weber, S., Li, Y., Seaman III, J.W., Kakizume, T. and Schmidli, H. (2021) Applying Meta-Analytic-Predictive Priors with the R Bayesian Evidence Synthesis Tools. *Journal of Statistical Software*, **100**, 1–32.

West, M. (1993) Approximating Posterior Distributions by Mixture. *Journal of the Royal Statistical Society. Series B (Methodological)*, **55**, 409–422.

NOVARTIS