EFSPI & BBS “Small populations and level of evidence”

Bayesian analysis for small sample size trials using informative priors derived from historical data

Andreas Kaiser
Bayer AG, Pharmaceuticals

27 Jun 2018

Introduction

Planning a trial in patients with ΔF508 mutation in cystic fibrosis at early clinical development

- Cystic fibrosis is a multi-organ disorder
  - Caused by gen mutations, cystic fibrosis transmembrane conductance regulator (CFTR) channels are affected; most frequent mutation is ΔF508
  - Today survival depends on status of lung disease, median age of survival is 40 years
  - Rare disease: Affects ≈1 out of 3000 newborns of Northern European ancestry, ≈30000 patients with CF in USA
- At early clinical development, trials are conducted to evaluate the treatment potential of drug
  - Desire to conduct trials with small sample sizes
    - Large challenges in recruiting patients for trials in rare disease like cystic fibrosis
    - Limit number of patients to be exposed to drug without proven clinical benefit
    - Fast decision making to avoid delaying development of potential efficacious drug
  - Utilize historical information by using Bayesian approach
    - Replace real patients by “virtual” patients → reduce sample size

Trial design and historical data

Early clinical development trial in patients with ΔF508 mutation in cystic fibrosis

- Randomized, placebo-controlled, parallel group trial in cystic fibrosis patients with by ΔF508 mutation
- Primary variable: Change of sweat chloride (Cl) content from baseline
  - Sweat Cl content is established diagnostic and biomarker for clinical trials for cystic fibrosis
- Primary objective: Evaluate reduction of sweat Cl content from baseline under treatment vs. placebo
- Approach: Use historical data to reduce number of patient with ΔF508 mutation for placebo treatment
  - Historical data for change from baseline of sweat Cl content (mmol/L) in patients treated with placebo

Source	Gender	Age	FEV1 pred.	Genotype	Time	N	Mean	St. dev.
Clancy et.al., Thorax (2012)	F & M	18 - 54	>= 40%	F508del-CFTR, homozygote	Day 14	17	1.75	7.7
Flume et.al., Chest (2012)	F & M	12 - 52	>= 40%	F508del-CFTR, homozygote	Day 15	26	-0.04	8.1
Boyle et.al., Lancet (2014)	F & M	>= 18	>= 40%	F508del-CFTR, homozygote	Day 14	21	-1.7	8.7
Ramsey et.al., N Engl J Med (2011)	F & M	12 - 53	31.6 - 98.2	G551D-CFTR, homozygote	Day 15	74	-0.15	7.8

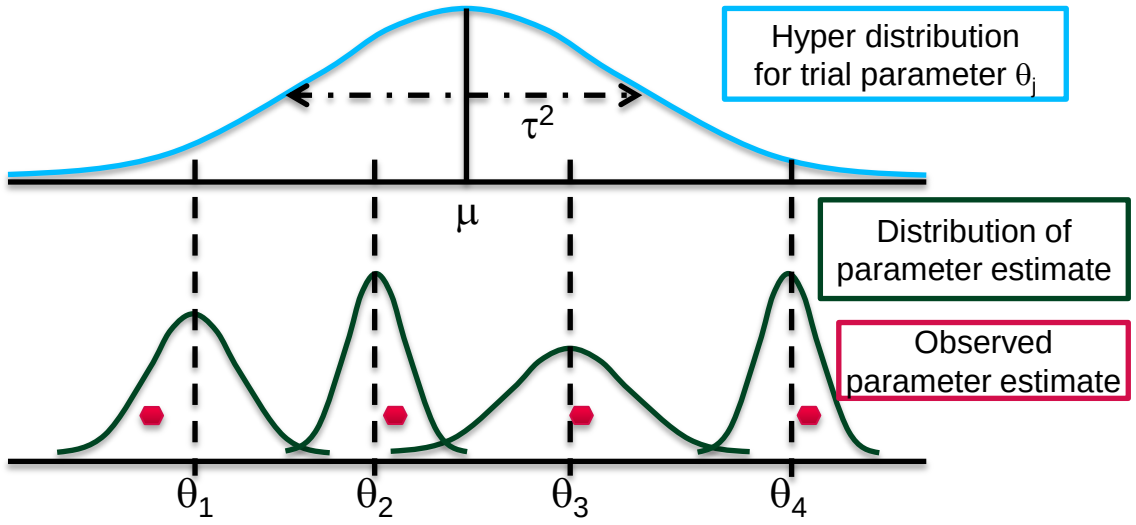
  - Selection of data based on external expert feedback (status of 2015)
    - Similar changes for both genotypes for placebo
    - Inclusion of recent articles to ensure comparable measurement procedure for sweat Cl
  - Between trial-variability for mean and potentially for standard deviation

Deriving prior distribution from historical data sources

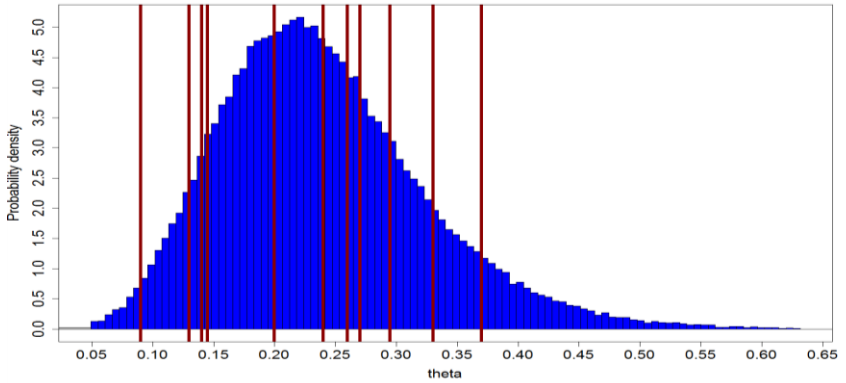
Using meta-analytical prediction to deal with between-trial variability

- Deriving prior distribution for placebo by simple pooling of historical data ignores between-trial variability
  - Would lead to overestimation of existing information
- Meta-analytical prediction (MAP) approach allows deriving of prior distribution while taking uncertainty due to between-trial variability into account [Neuenschwander et al. (2010), Schmidli et al. (2014)]

1. Bayesian random-effect meta analysis on historical data for trial parameter, e.g. means



2. Predicted posterior distribution for parameter θ^* in new trial



Predicted posterior distribution
= Hyper distribution + uncertainty in estimates of μ and τ^2

- Predicted posterior distribution for parameter θ^* in new trial reflects all information about this parameter

Modelling estimates for mean and variance for placebo in trials

MAP: Model for Bayesian random-effect meta analysis for sweat Cl content

- Response $y_{ij},\ i = 1, \ldots, n_j$ of patient *i* in trial *j* is normal distributed with unknown, trial-specific mean ϑ_j and variance σ_j^2

$$y_{ij}|\vartheta_j, \sigma_j^2 \sim N\left(\vartheta_j, \sigma_j^2\right)$$

  - Observed mean response M_j of trial *j* follows a normal distribution

$$M_j|\vartheta_j, \sigma_j^2 \sim N\left(\vartheta_j, \frac{\sigma_j^2}{n_j}\right)$$

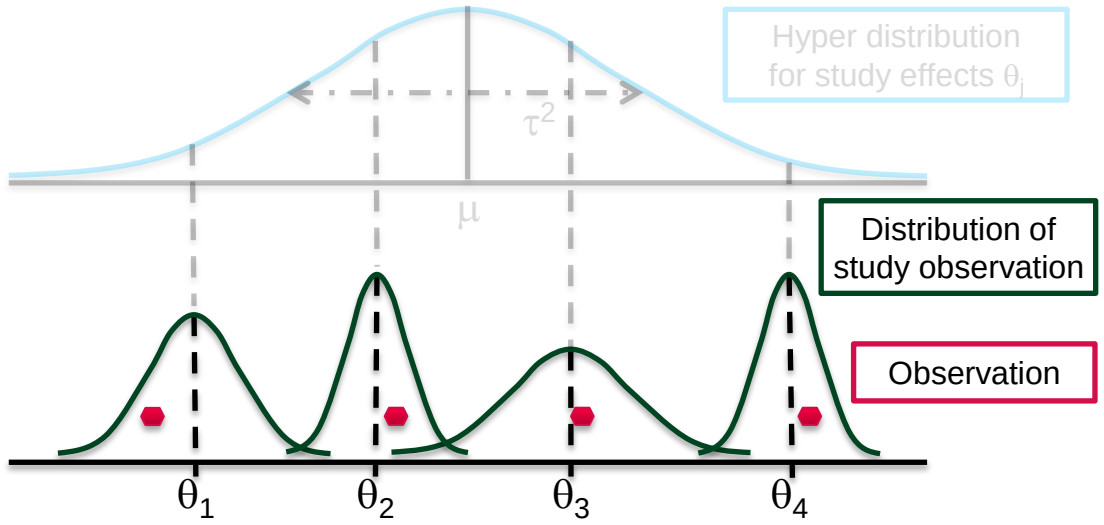
  - Observed variance of response S_j^2 of trial *j* follows a "scaled" χ^2-distribution

$$S_j^2|\ \sigma_j^2 \sim \chi^2(n_j - 1, \sigma_j^2)$$

    i.e. $S_j^2/\sigma_j^2 \sim \chi^2(n_j - 1)$ follows a χ^2-distribution with $n_j - 1$ degree of freedom

Bayesian random-effect meta analysis on historical data

- Modelling of observations (e.g. trial means)



Modelling the unknown means of sweat Cl content in trials

MAP: Model for Bayesian random-effect meta analysis for sweat Cl content

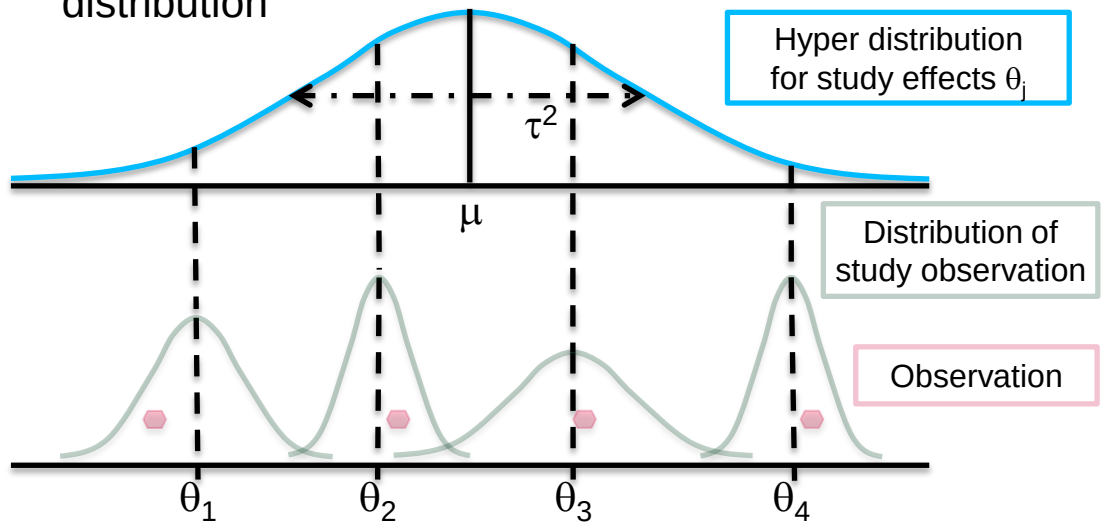
- For modelling the means ϑ_j of the trials, a normal distribution is used with unknown mean μ and variance τ^2 as hyper distribution

$$\vartheta_j | \mu, \tau^2 \sim N(\mu, \tau^2)$$

- Weak informative prior distributions for hyper distribution parameters are used
  - Mean μ ~ Normal distribution
    - mean = 0, variance = 5^2
  - Variance τ^2 ~ Gamma distribution
    - mean = 1.5^2, CV = 100%
  - Use of informative priors to cope with small numbers of historical trials [Wandel et al. (2017), Friede et al. (2017)]
    - Values were chosen to restrict μ and τ^2 on plausible values while avoiding “over domination”
    - Evaluation of impact of informative priors in meta-analysis by sensitivity analysis

Bayesian random-effect meta analysis on historical data

- Modelling trial parameters (e.g. means) by hyper distribution



Modelling the unknown variance of sweat Cl content in trials

MAP: Model for Bayesian random-effect meta analysis for sweat Cl content

- For modelling the variance σ_j^2 of the trials, an inverse-gamma distribution is used as hyper distribution
  - To allow easier assessing of between-trial variability of variance, the inverse gamma distribution is reparametrized by the mean δ^2 and the coefficient of variation ε

$$\sigma_j^2|\delta^2,\varepsilon \sim \text{Inv}-\Gamma(2+1/\varepsilon^2, \delta^2\,(1+1/\varepsilon^2))$$

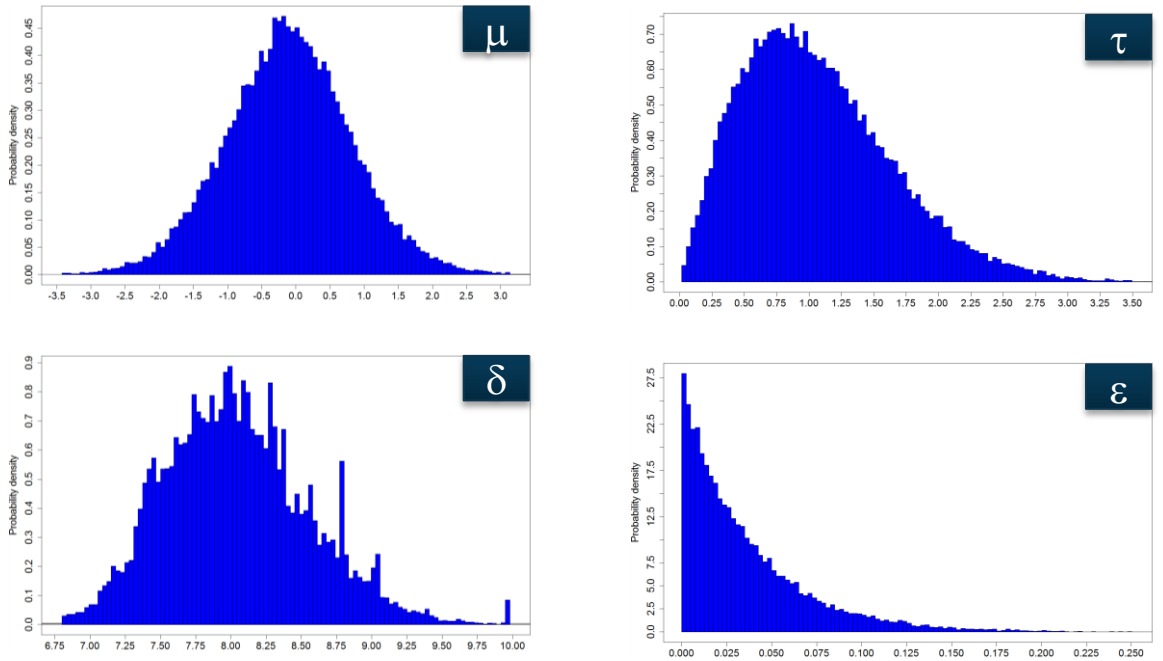
- Weak informative prior distributions for hyper distribution parameters are used
  - Mean δ^2~ Gamma distribution
    - mean = 8^2, CV = 100%
  - Coefficient of variation ε~ Gamma distribution
    - mean = 20%, CV = 100%

- Remark: For common variance, $\sigma^2 = \sigma_j^2$, replace hyper-distribution by inverse-gamma distribution for σ^2 as prior distribution
  - Affects mainly variance estimates of studies, small impact on marginal posterior for trial means

Results of Bayesian meta-analysis for sweat Cl content for placebo

Results of first step of meta-analytical prediction

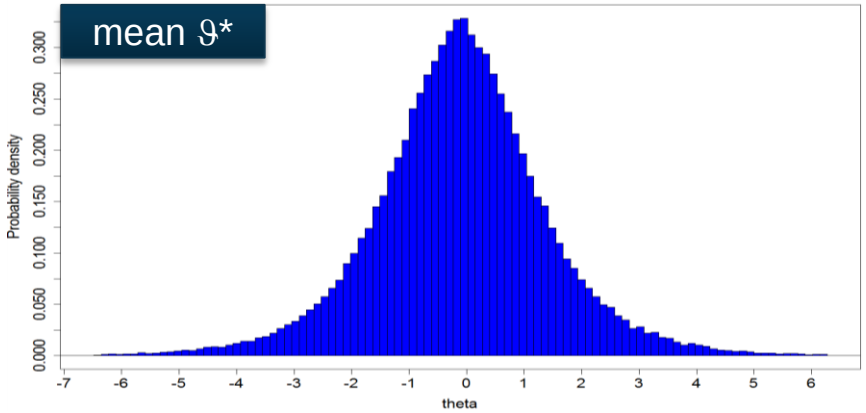
- Posterior distribution for parameter of hyper-distributions are derived by MCMC simulations
- Assessment of posterior probabilities provide insight about “reproducibility” of endpoints
  - Trial means ϑ_j
    - Located around $\mu \approx 0$
    - Small between-trial variability, $\tau \approx 1$
  - Variance of trials σ_j^2
    - Located around $\delta^2 \approx 8^2$
    - Negligible between-trial variability, $\varepsilon \approx 4\%$
  - Remark: Peaks in histogram for δ indicate convergence issues in MCMC simulation caused by limited information from data due to few studies → need for informative priors

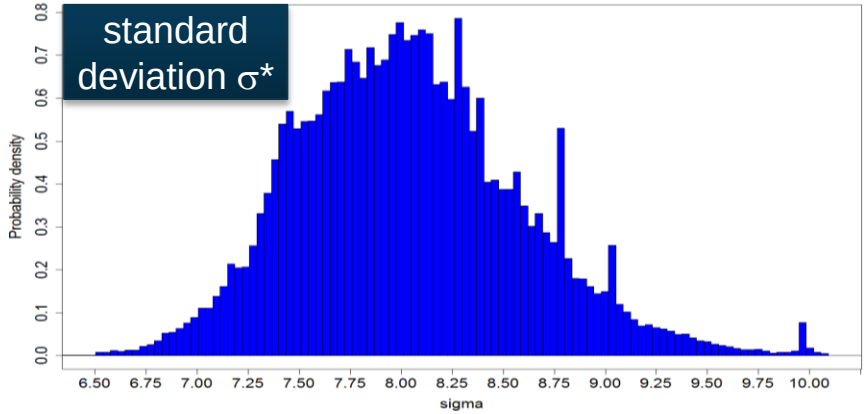


Information about mean and variance of future trials

Results of prediction step of meta-analytical prediction

- Knowledge about mean ϑ^* and variance σ^{*2} of a future trial is reflected by the hyper distributions taking uncertainties in for distribution parameters into account
  - Predicted posterior distribution for ϑ^* and σ^{*2} is a representation for the available information
  - MC samples for ϑ^* and σ^{*2} are derived from hyper distribution using MCMC parameter samples
- Marginal predicted posterior distribution for trial mean ϑ^* and standard deviation σ^*





Parameter	mean	sd	q1	median	q3	90%-range_low	90%-range_up
Mean of new trial ϑ*	-0.11	1.54	-0.98	-0.11	0.76	-2.58	2.37
Standard deviation of new trial σ*	8.04	0.56	7.65	8.01	8.38	7.21	9.02

Deriving prior distribution for sweat Cl content for Bayesian analysis

Use predicted posterior distribution for future trial from MAP to derive prior information

- The predicted posterior distribution for ϑ^* and σ^{*2} reflects all prior information for mean and variance
  - Use predicted posterior distribution as prior distribution for placebo
- Derive analytical form for prior distribution by fitting a Normal – Inv-χ^2 distribution to MC samples of predicted posterior distribution

$$\vartheta_{placebo}, {\sigma_{placebo}}^2 \sim \text{Normal} - \text{Inv}{-}\chi^2(\text{mean} = -0.1, \text{kappa} = 27.8, \text{dof} = 106.3, \text{sd}^2 = 8^2)$$

  - Trial mean for placebo: Amount of information corresponds to ≈28 patients (kappa = 27.8)
    - ≅ 20% of total sample size in historical trials (N=138)
      → information reduction caused by between-trial variability
  - Trial variance for placebo: Amount of information corresponds to ≈107 patients (dof = 106.3)
    - ≅ 80% of total sample size
      → small information reduction due to negligible between-trial variability
  - Remark: To cope with long tails and/or shape deviations from standard distribution, deflate prior distribution (information loss) or use approximation by mixture distribution [Schmidli et al. (2014)]

Evaluation of treatment potential of drug using Bayesian approach

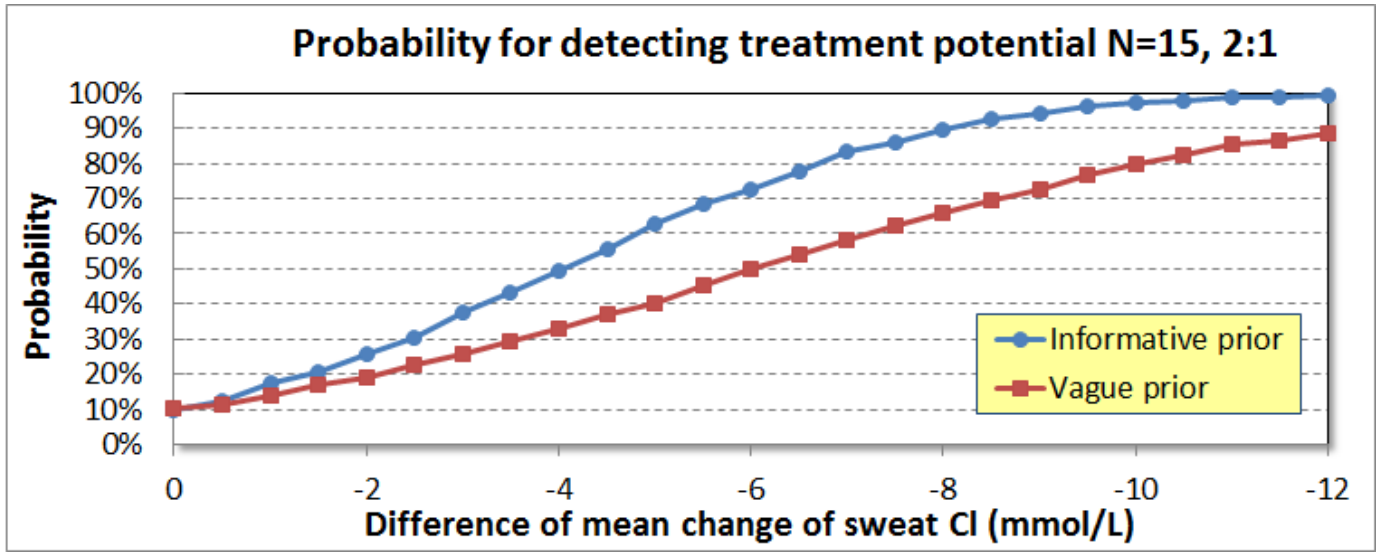
- Assessment of treatment potential by Bayesian analysis: Evaluate difference in mean change of sweat Cl content from baseline between drug and placebo $\vartheta_{drug} - \vartheta_{placbo}$

  - Marginal posterior probability for mean change of sweat Cl from baseline for placebo ϑ_{placbo}
    - Incorporate prior information for placebo derived by MAP procedure
      - Based on expert feedback, no further deflation of prior despite potential domination over data
  - Marginal posterior probability for mean change of sweat Cl from baseline for drug ϑ_{drug}
    - Vague prior information
  - Posterior distribution for mean difference $\vartheta_{drug} - \vartheta_{placbo}$
    - MC sampling from marginal posterior distributions for ϑ_{placbo} and ϑ_{drug}
    - Take differences of MC samples for drug and placebo

Operational characteristic of assessment of treatment potential

- Capability of planned trial to detect treatment potential is evaluated by trial simulations
  - Assess probability for detecting treatment potential for fixed difference in means $\vartheta_{drug} - \vartheta_{placbo}$
  - Coping with uncertainty about ϑ_{placbo} and ${\sigma_{placbo}}^2$: Sample parameters and averaging probability using prior distribution derived by MAP [Walley et.al. (2015)]; for drug higher variability is assumed $\sigma_{drug} = 1.1 \cdot \sigma_{placbo}$

- Sample size determination
  - 2:1 randomization
  - $n_{drug} = 10$ and $n_{placebo} = 5$
  - $\vartheta_{drug} - \vartheta_{placbo}$ = -8 mmol/L
    → ≈ 90% probability for detecting treatment potential on average



- Incorporation of prior information increases probability to detect treatment potential
  - Allows designing trial with smaller sample sizes and unbalanced randomization

Summary and conclusions

- Trials with small sample sizes are desired in rare disease, in particular at early stage of development
- Incorporation of historical information might allow reduction of patient numbers
  - Bayesian approach provides formal framework for incorporation of prior information in trial analysis
  - Relevant sample size reduction can be achieved for small trials allowing imbalanced designs
- Selection of historical data is a crucial step when deriving prior information for Bayesian analysis
  - Close interaction with clinicians and experts is essential
  - Meta-analytical prediction is a useful approach to derive prior distributions from historical source
    - Cope with between-trial variability, thus prevents overestimation of information
      - Usually effective sample size of derived prior distribution is much smaller than total sample size
    - Provides additional insight about e.g. mean and variance of treatment response
    - Meta-analysis on few historical trials is feasible when incorporating (weak) informative priors
      - Evaluate impact of incorporating informative prior by sensitivity analysis
  - Domination of prior information in analysis and deflation of prior distribution needs to be considered

Thank you!

Bye-Bye