



MRC  
Biostatistics  
Unit



UNIVERSITY OF  
CAMBRIDGE

---

## **Bayesian Methods for Clinical Trials**

*Lecture 8: Design and analysis of clinical trials using historical information*

Libby Daniels & Pavel Mozgunov & Thomas Jaki

MRC Biostatistics Unit

Feb 2025

# Acknowledgments

- Dr Gaëlle Saint-Hilary (Saryga SAS, France)

# Introduction

- In conventional trials, the analysis is based only on **current patients**
- Usually, historical information is available on the control arm and/or treatment effect, but it is **not formally used**
- There is a growing interest in increasing the efficiency of new clinical trials by using this **historical/external information**
  - ▶ Potential of reducing the number of patients, which can reduce the cost of the trial and make trial more attractive for patients
  - ▶ Considering the whole amount of evidence also permits to improve decision-making in drug development

# Introduction

- There is an increasing **methodological comfort**, even for confirmatory trials
- Regulators are increasingly open to the use of historical information borrowing
- The main challenge is the selection of appropriate historical information
- There are both **benefits and risks** to historical information borrowing

*Risk of false positive conclusion, power, bias, precisions considerations*

# Some available methods

Pocock's method

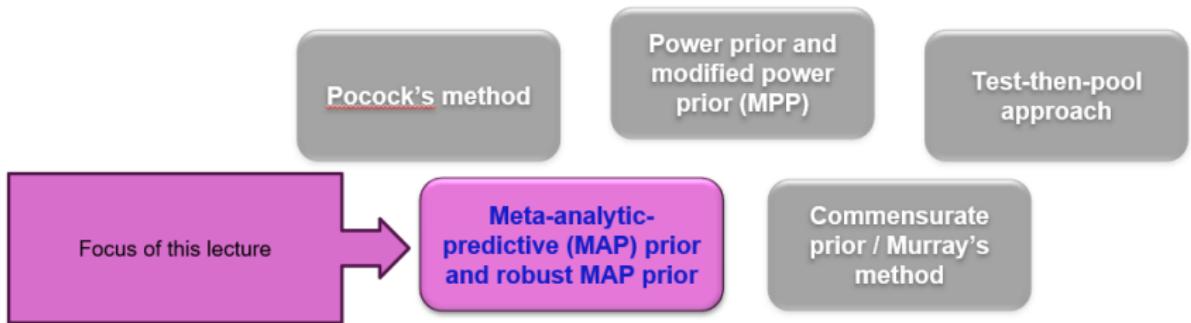
Power prior and  
modified power  
prior (MPP)

Test-then-pool  
approach

Meta-analytic-  
predictive (MAP) prior  
and robust MAP prior

Commensurate  
prior / Murray's  
method

# Some available methods



# Motivating example

Double-blind, randomised PoC study

- Secukinumab vs placebo (2:1 ratio)
- Moderate to severe Crohn's disease
- **Primary endpoint:** Crohn's Disease Activity Index (CDAI) change from baseline at week 6
- **Historical placebo data:** 6 studies (summary statistics) with 671 patients in total

*Q: Can we use available evidence to estimate the treatment effect with precision while reducing the number of patients on placebo?*

# Meta-Analytic Prior (MAP) Approach

- **Historical placebo data** are modelled via **Hierarchical model**  
Parameters  $\theta_h$  for the placebo effect in the study  $h = 1, \dots, H$

$$y_h | \theta_h \sim N(\theta_h, s_h^2)$$

- **Current study:** Parameter  $\theta_*$  for the placebo effect
- **MAP Prior:** Predictive distribution of  $\theta_*$  based on historical evidence only i.e distribution of  $\theta_* | y_1, \dots, y_H$  derived from

$$\theta_1, \theta_2, \dots, \theta_H, \theta_* \sim N(\beta, \tau^2)$$

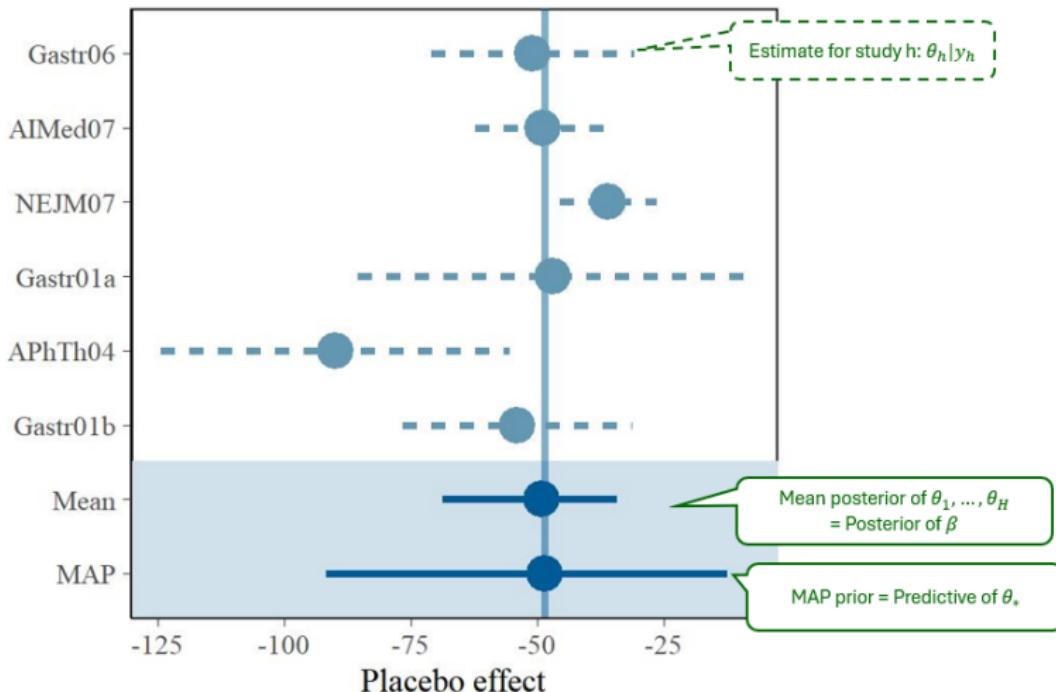
$$\text{Normal prior for } \beta \sim N(0, \sigma^2)$$

$$\text{Half - Normal prior for } \tau \sim HN(0, a)$$

$\sigma$  is within-trial standard deviation of the endpoint

$\tau$  is between-trial standard deviation

# MAP Illustration with RBesT



# Prior Effective Sample Size

- **Prior Effective Sample Size (ESS)** is quantity of information introduced through the MAP prior expressed in terms of number of patients
- This is not always straightforward
  - ▶ It is not simply the number of historical control subjects
  - ▶ It also depends on the between-trial heterogeneity ( $\tau$ )

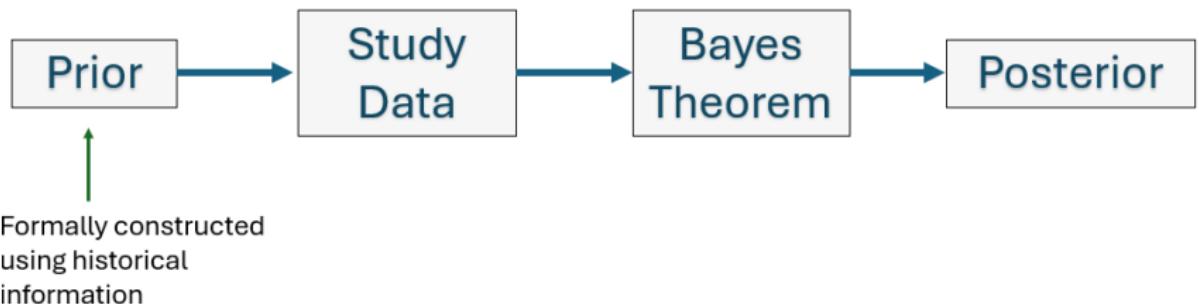
# Prior Effective Sample Size Estimation

- The **moment's method** calculates the ESS of a conjugate distribution with the same mean and SD as the MAP prior
- Methods Based on Fisher's information
  - ▶ **Morita et al. (2008)** locally evaluate the Fisher information at the mean, the mode or the median of the distribution of  $\theta$   
*May lead to optimistic ESS values (i.e. large)*
  - ▶ **Neuenschwander et al. (2019)** – evaluate the mean ESS minimizing the distribution of Fisher Information (ELIR method)

Prior Maximum Effective Sample Size is  $\frac{\sigma^2}{\tau^2}$

Even if historical information is large, **the contribution to the analysis in the next trial could be small** if the between-trial heterogeneity is large

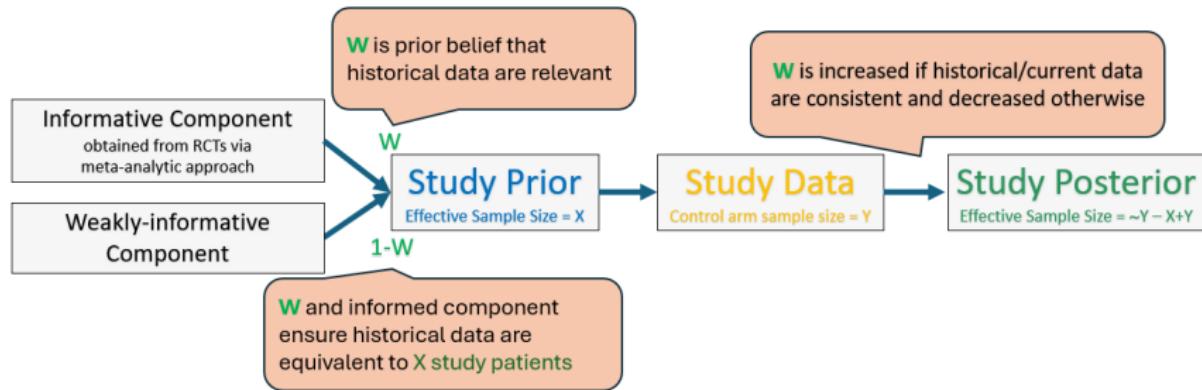
# Bayesian borrowing



# Bayesian Dynamic Borrowing (BDB)

- Historical study data are used to construct the prior distribution
- If the historical and current study data **are not similar**, Dynamic Borrowing allows to **account for the inconsistency** between these by adjusting how much information to borrow;
- The larger the inconsistency, the less we borrow
- The smaller the inconsistency, the more we borrow

# Bayesian Dynamic Borrowing (BDB)



# Robustification of the MAP prior

- **Robustness against prior-data conflict**
- Mixture prior with two components Robustness against prior-data conflict  
*(Informative) MAP prior and (Weakly informative) vague prior*

$$f(\theta_*|y_1, \dots, y_H) = w_R \times f_V(\theta) + (1 - w_R) \times f_{MAP}(\theta_*|y_1, \dots, y_H)$$

where  $f_V(\cdot)$  is the density of a vague prior

# How does a robust prior protect against prior-data conflicts?

- $w_R$  is prior probability that historical data **are not relevant**
- $(1 - w_R)$  is prior probability that historical data **are relevant**
- For mixture prior distribution, the posterior distribution is a mixture distribution with **updated weights**
- Let  $p_R$  be likelihood of new data  $y_*$  generated from vague prior
- $p_{MAP}$  likelihood of new data  $y_*$  generated from MAP prior

$$w'_R = \frac{w_R \times p_R}{w_R \times p_R + (1 - w_R) \times p_{MAP}}$$

Once we get the new data, the posterior distribution of  $\theta$  gives:

- ▶ More weight (than prior) to the vague if data more aligned with it
- ▶ More weight (than prior) to MAP prior if data more aligned with it

*This is called dynamic borrowing*

## How to choose $w_R$ ?

- Reflects the possibility that the control group in the current trial is different to the historical control group population
- Often interpreted as the **prior level of confidence** that MAP component is “most suitable” than the vague one
- For the **unit-information prior**, the usual choice is 10% or 20%
- There is an **interplay** between the weight and effective sample size of the vague component (should be chosen together)

$$w'_R = \frac{w_R \times p_R}{w_R \times p_R + (1 - w_R) \times p_{MAP}}$$

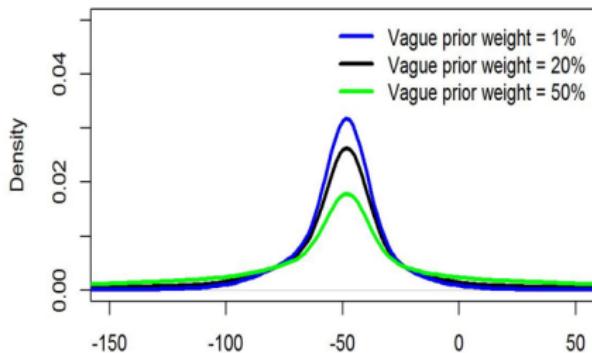
- **Operating characteristics** are essential to choose the weight and the ESS of the vague component

# rMAP in the motivating trial

More weight on the vague prior increases the variance of the robust MAP prior

More weight on the vague prior decreases the ESS

Robust MAP priors



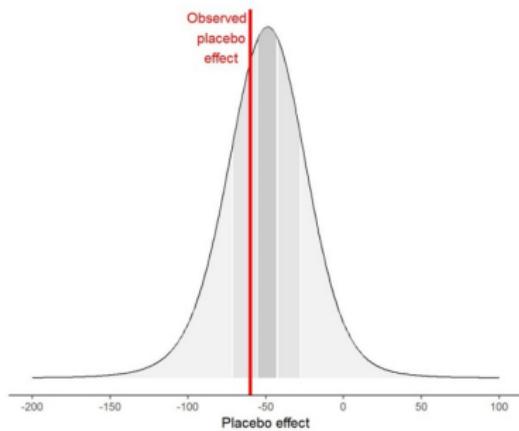
Vague prior weight	ESS (ELIR method)
0% (no robust)	37.3
1%	36.7
20%	27.3
50%	14.5

# Analysis in the motivating trial

<b>Analysis</b>	<b>TE</b>	<b>Active</b>	<b>Placebo</b>
Primary	-45.6 (-82, -6)	-98.8 (-126, -72)	-53.1 (-83, -30)
Current only	-40.0 (-88, 8)	-100.0 (-128, -72)	-60.0 (-99.6, -20.4)
MAP			-49.7 (-93, -9)
rMAP			-49.7 (-152, 52)

# Analysis in the motivating trial

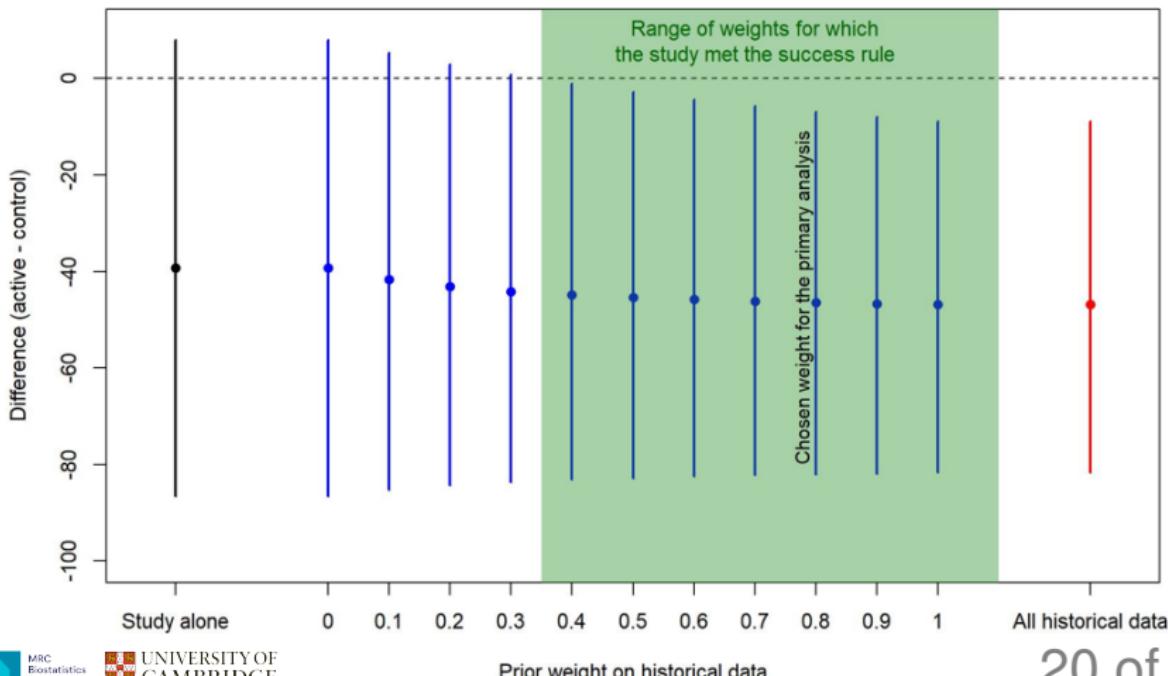
**Prior predictive distribution:** predicted distribution of observed placebo effect in new study



- Consistency between historical and new studies
- Reflected by prior and posterior weights:
  - ▶ Prior weight  $W = 0.80$
  - ▶ Posterior weight  $W = 0.93$

# Tipping Point Analysis

Posterior median and 95% CrI for the estimated treatment difference vs prior weight on historical placebo data



# Prospective planning of a trial with historical information borrowing

- Bayesian probability of success

$$\text{Prob}[TE \geq 0 \mid \text{Data, Prior}] > c$$

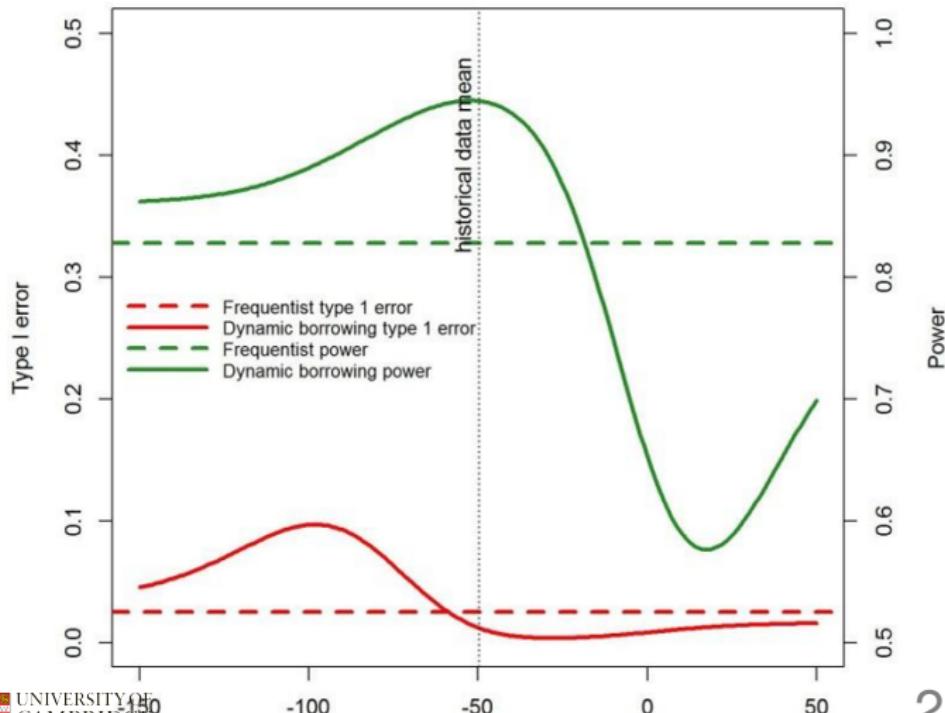
- Type I Error: calculated under  $TE = 0$
- Power: calculated under  $TE = TE^*$

However, the TE is not the only assumption that we have to make when evaluating the operating characteristics of BDB.

**The extent of the consistency of the historical and current data**

# Operating characteristics

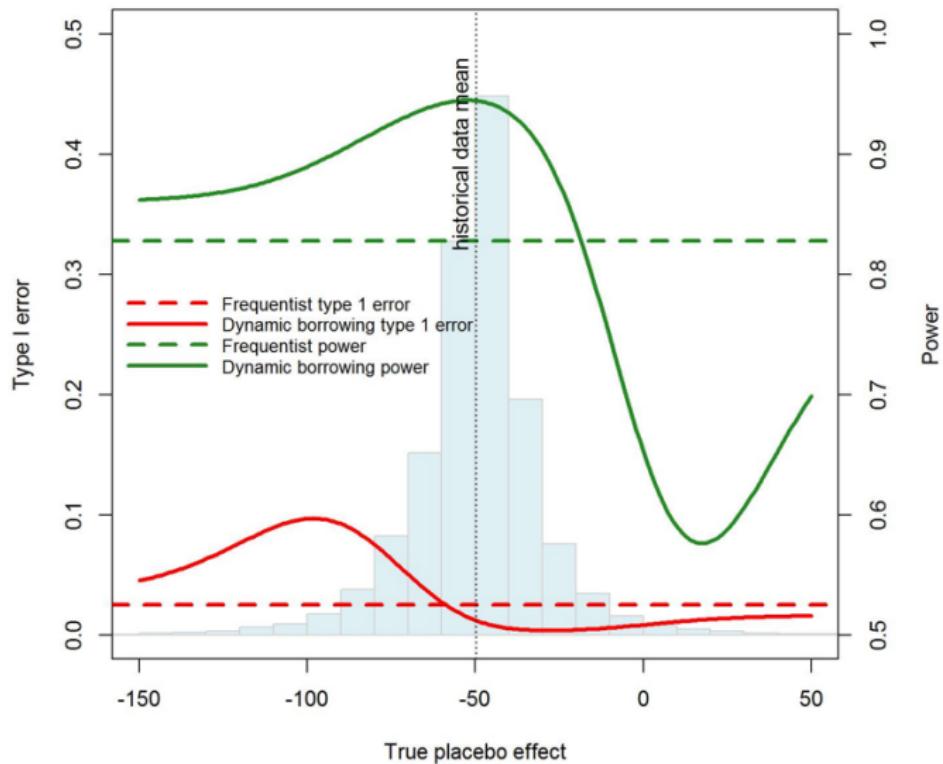
- **Type 1 error:** Probability of false positive given TE = 0
- **Power:** Probability of true positive result given TE = -70



# Advanced Operating Characteristics

- **Analysis Prior** – used in analysis of the current trial and represents best reflection of the evidence and the corresponding uncertainty
- **Design Prior** – used for design evaluation to calibrate Bayesian designs under different assumptions about the true parameter value(s)

# Advanced Operating Characteristics



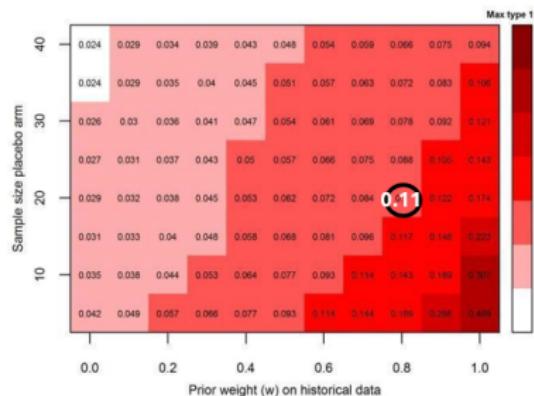
# Advanced Operating Characteristics

- Pointwise type 1 error (no drift) = 1.3%
- Max pointwise type 1 error = 11.0%
- Average type 1 error, over MAP prior = 2.2%
- Average type 1 error, over robust MAP prior = 2.5%

# Choosing design parameters

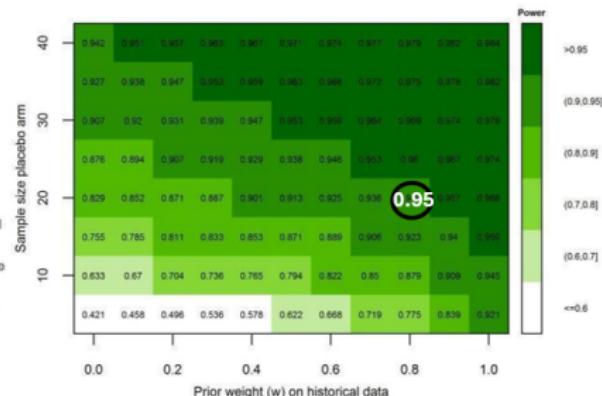
Heatmaps to understand the impact of different assumptions on the risk of false positive and false negative conclusions

Maximum type I error



Power

Assuming consistency with historical data



# Further metrics

Assumed observed placebo mean response value in new study	Proposed BDB design ( $w = 80\%$ )		Frequentist design
	Posterior weight ( $\tilde{w}$ ) on historically-informed evidence	Point estimate (posterior mean) of the placebo response [95% CrI]	Point estimate of the placebo response [95% CI]
-20	0.88	-36.6 [-56.1; -8.9]	-20 [-58.6; 18.6]
-50	<b>0.93</b>	-48.9 [-69.7; -30.4]	-50 [-88.6; -11.4]
-80	0.87	-63.5 [-93.2; -41.2]	-80 [-118.6; -41.4]
-110	0.56	-93.8 [-132.0; -57.3]	-110 [-148.6; -71.4]
-140	<b>0.19</b>	-132.6 [-165.9; -96.2]	-140 [-178.6; -101.4]

# Conclusions

- Use of historical information permits to **increase the level of confidence** in the decisions or maintain the level of confidence and reduce the number of control subjects in the studies
- However, it can lead to the **type I error inflation** in case of prior-data conflict
- **Using dynamic borrowing** limits type I error inflation in case of prior-data conflict
- But also reduces the amount of borrowing in case of no conflict (compared to MAP)
- There is no free lunch...(Kopp-Schneider et.al 2020)

# Main references

- Best N, Ajimi M, Neuenschwander B, Saint-Hilary G, Wandel S. Beyond the classical Type I error: Bayesian metrics for Bayesian designs using informative priors. *Stat Biopharm Res.* 2024 May 31. doi: 10.1080/19466315.2024.2342817.
- Hueber F, Secukinumab in Crohn's Disease Study Group. 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.* 2012 Dec;61(12):1693–1700. doi: 10.1136/gutjnl-2011-301668.
- Kopp-Schneider A, Calderazzo S, Wiesenfarth M. Power gains by using external information in clinical trials are typically not possible when requiring strict type I error control. *Biom J.* 2020;62:361–374. doi: 10.1002/bimj.201800395.
- Morita S, Thall PF, Müller P. Determining the effective sample size of a parametric prior. *Biometrics.* 2008;64:595–602. doi: 10.1111/j.1541-0420.2007.00888.x.
- Neuenschwander B, Weber S, Schmidli H, O'Hagan A. Predictively Consistent Prior Effective Sample Sizes. *Biometrics,* 2019
- Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan O, Spiegelhalter DJ and Neuenschwander B. Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 2014; 70: 1023-32.
- Viele K, Berry S, Neuenschwander B, et al. Use of historical control data for assessing treatment effects in clinical trials. *Pharm Stat.* 2014;13:41–54. doi: 10.1002/pst.1589.
- Weber S, Neuenschwander B, Schmidli H, et al. RBesT: R Bayesian Evidence Synthesis Tools. R package version 1.6-1. Available at: <https://cran.r-project.org/web/packages/RBesT/index.html>.