Using Elastic Prior to Design Clinical Trials with Adaptive Information Borrowing

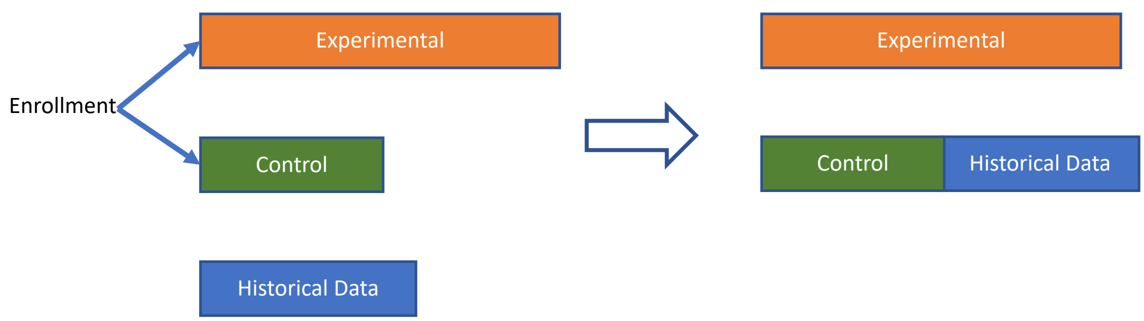
Ying Yuan
MD Anderson Cancer Center

Liyun Jian (MDACC) and Lei Nie (FDA)

Outline

- Introduction of challenges and methods for borrowing information from historical data
- Elastic prior approach
- Conclusion

Borrowing Information from Historical Data



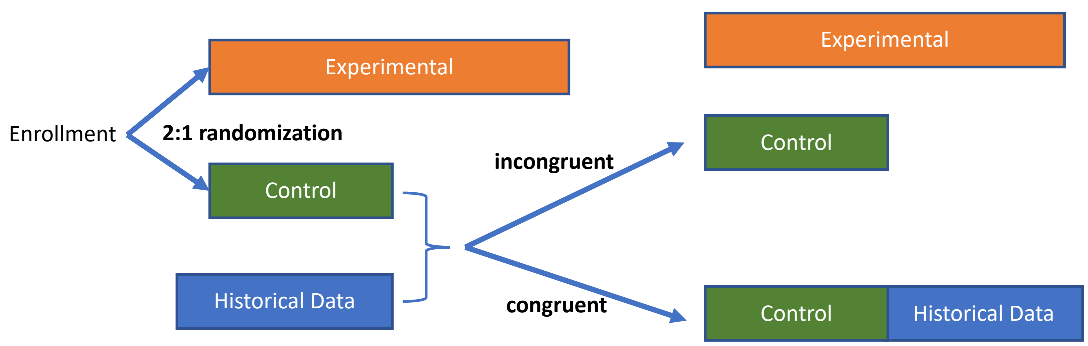
A Double-edged Sword

- Historical data may or may not be congruent to trial data due to various reasons
  - Different patient population, eligibility criteria, treatment procedure, facilities, care providers…

Sc.	Historical Mean	Control Mean	Experimental Mean	Consequence of borrowing
A	1	1	2	Increase power/reduce sample size
B	0.5	1	1	Inflated type I error
C	1.5	1	2	Reduced power

Adaptive Information Borrowing

- Borrow if historical data are congruent to the control (of the trial); otherwise, do not borrow



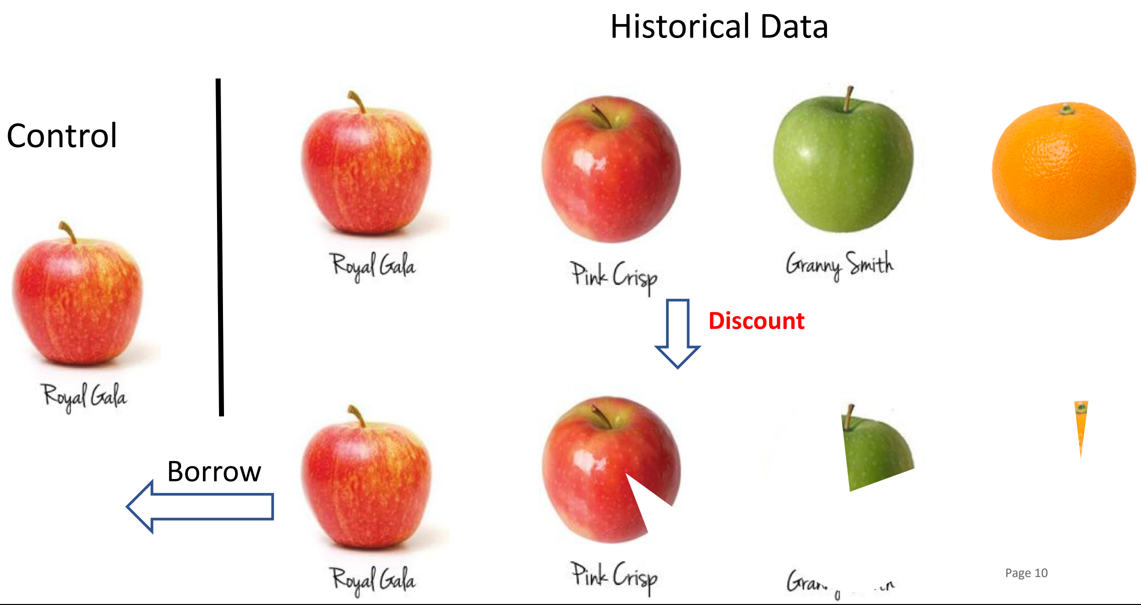
Page 7

Bayesian Methods for Information Borrowing

- Power Prior (Ibrahim and Chen, 2000)
- Commensurate prior (Hobbs, et al., 2011)
- Robust meta-analysis-predictive prior (Schmidli, et al., 2014)
- Bayesian hierarchical model (Bernardo and Smith, 1994)

Page 9

Power Prior



Power Prior

- Let θ denote the treatment effect
- Let $\pi_0(\theta)$ denote the prior distribution of θ (before accounting for the historical data), typically specified as noninformative prior
- Let D_0 denote the historical data, and D denote the trial data
- Power prior is given by

$$\pi(\theta | D_0, \delta) \propto L(D_0|\theta)^{\delta} \pi_0(\theta)$$

  - $L(D_0|\theta)$ is the likelihood of historical data
  - δ is the **power parameter**, controlling how much information to be borrowed from D_0

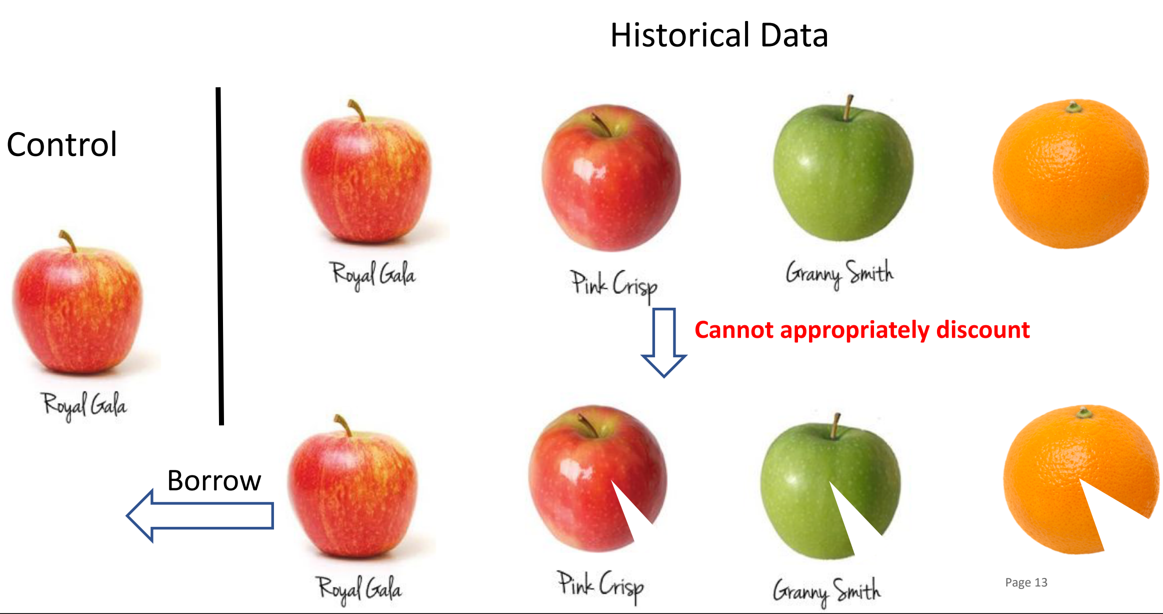
Power Prior

- Adding the trial data D, the posterior of θ arises as

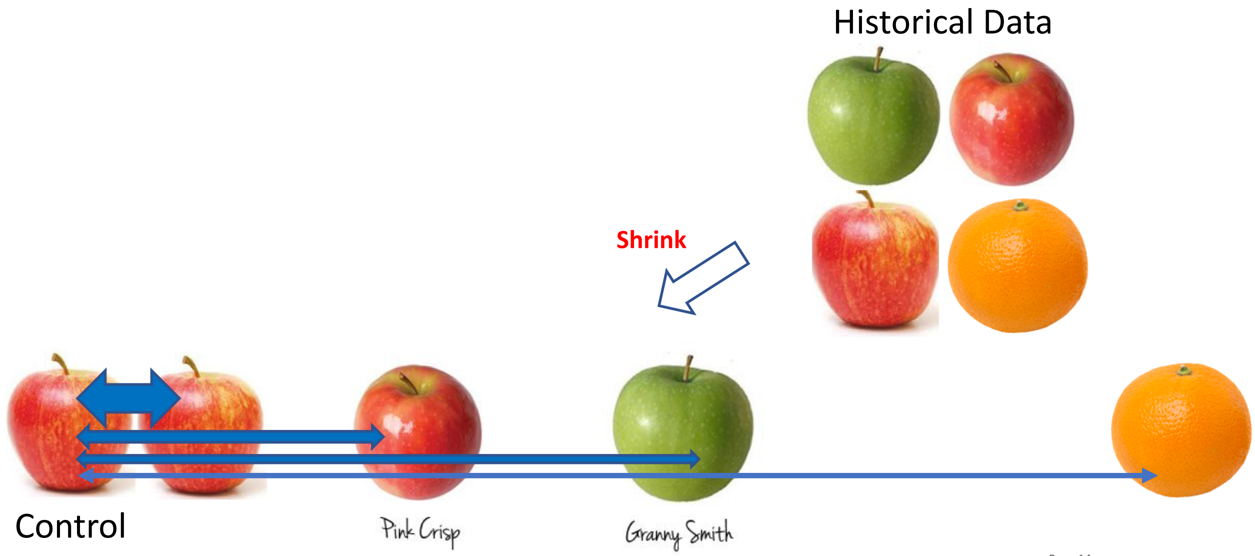
$$\pi(\theta|D, D_0, \delta) \propto \pi(\theta|D_0, \delta) L(D|\theta)$$
$$= \pi_0(\theta) L(D_0|\theta)^{\delta} L(D|\theta)$$

- When $\delta = 1$, fully information borrowing from D_0, and when $\delta = 0$, no information borrowing
- **Key issue**: how to determine the value of δ ?
- A natural approach is to assign δ a prior (e.g., uniform prior) and let data determine the degree of borrowing
- This approach, however, does not work well and leads to substantial type I error inflation (Neuenschwander et al., 2000; Pan and Yuan, 2017)

Difficulty of Power Prior



Commensurate Prior



Commensurate Prior

- Let θ_0 denote the true treatment effect in the historical data
- Commensurate prior is given by

$$\pi(\theta|D_0,\theta_0,\tau) \propto L(D_0|\theta_0)\pi(\theta|\theta_0,\tau)\pi_0(\theta)$$

- τ is the **shrinkage parameter**, controlling how much information to be borrowed from D_0
- Suppose $\pi(\theta|\theta_0,\tau) = N(\theta_0,\tau^{-1})$, then when $\tau \to 0$, little information borrowing, and when $\tau \to \infty$, full information borrowing

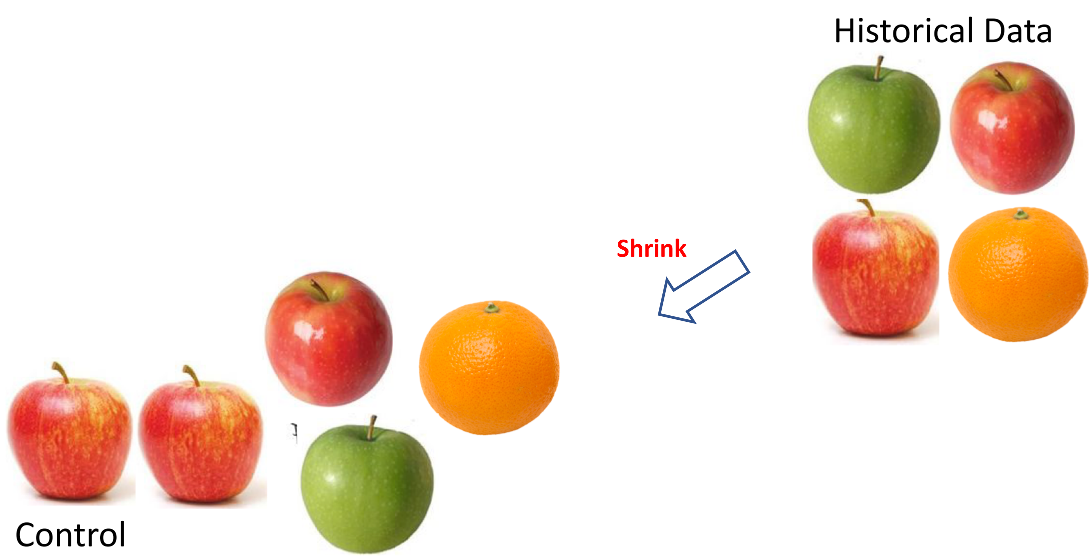
Commensurate Prior

- Given trial data D, the posterior of θ is

$$\pi(\theta|D, D_0, \tau) \propto \pi(\theta|D_0, \theta_0, \tau)L(D|\theta)$$
$$= \pi_0(\theta)\ L(D_0|\theta_0)\ L(D|\theta)\ \pi(\theta|\theta_0, \tau)$$

- Commensurate prior can be viewed as a special form of Bayesian hierarchical model geared to borrow from one historical dataset
- **Key issue**: how to determine the value of τ ?
- A natural approach is to assign τ a prior (e.g., uniform or spike-slab prior) and let data determine the degree of borrowing
- Similar issue as the power prior: cannot control information borrowing appropriately and lead to inflated type I errors

Difficulty of Commensurate Prior



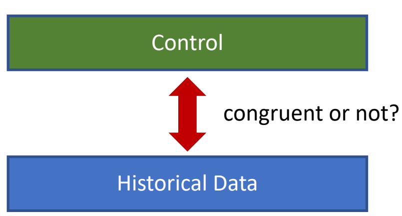
Simulation

N=50 for historical, 25 for control, 50 for experimental

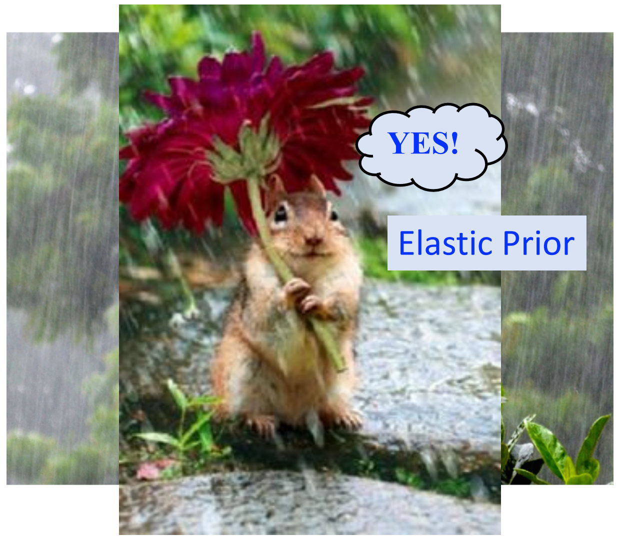
Mean			Type I error/Power (%)		
Historical	Control	Experimental	**No Borrow**	**Commensurate prior**	**Power Prior**
1	1	1	5.0	5.0	5.0
1	1	1.5	66.3	91.6	88.3

Intrinsic Difficulty for Adaptive Borrowing

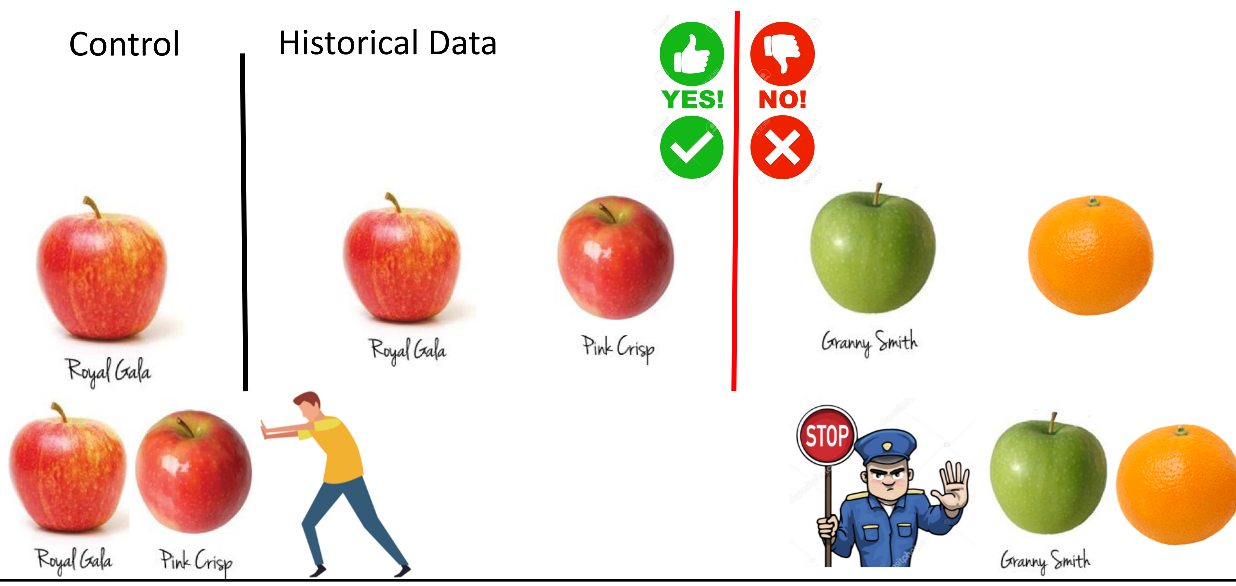
- Data contain very limited information to quantify the congruence/variation between the trial data and historical data
- Information unit is the number of datasets, not the number of subjects!



Can We Make Progress?



Elastic Prior



Elastic Prior

- Bayesian learning

Prior + Historical Data → Posterior

$$\pi_0(\theta)L(D_0|\theta) \to \pi(\theta|D_0)$$

→ New Prior + New Data → New Posterior

$$\pi(\theta|D_0)L(D|\theta) \to \pi(\theta|D_0, D)$$

→ ......

- Idea: discount the information in $\pi(\theta|D_0)$ by inflating its variance adaptively according to the elastic function

Page 22

Elastic Prior

- Suppose $\pi(\theta|D_0) = N(\theta_0, \tau^2)$
- Elastic prior is

$$\pi_E(\theta|D_0) = N(\theta_0, \tau^2/g(T))$$

where $g(T)$ is the elastic function, and *T* is a measure of congruence between D_0 and D.

Page 23

Elastic Function

- Consider a normal endpoint Y, let $\bar{y}_0$ and $\bar{y}$ denote sample means of D_0 and D, and S is standard error
- Define a metric T to measure the congruence between D_0 and D

$$T = \frac{|\bar{y} - \bar{y}_0|}{S^2}$$

- Elastic function is defined as

$$g(T) = \frac{1}{1 + \exp\{a + b[\log(T)]\}}$$

- To achieve adaptive information borrow, we enforce that $g(T)$ → 1 when when D_0 and D are congruent, and $g(T)$ → large (e.g., 100) when D_0 and D are incongruent.

Page 24

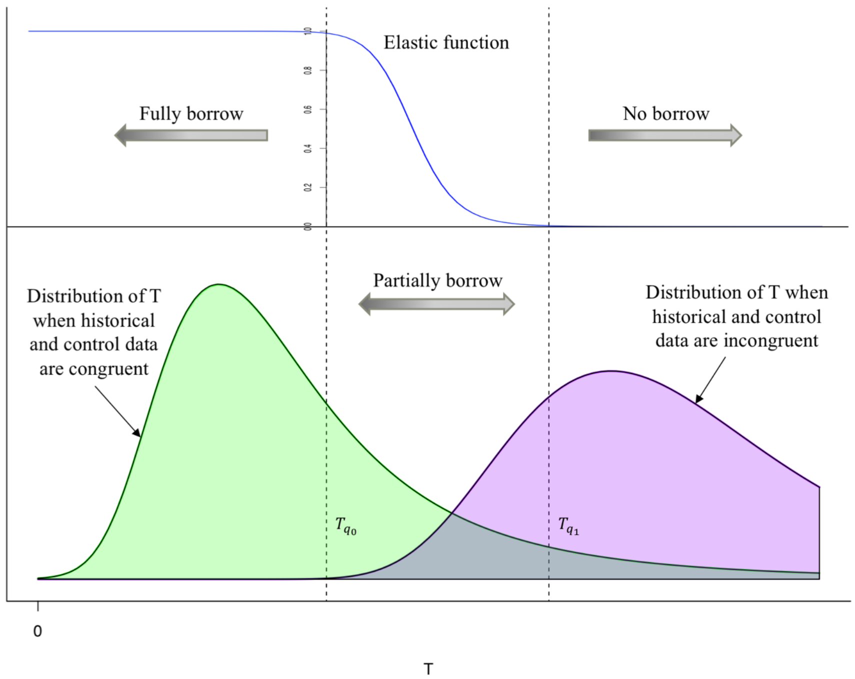
Determine Parameters of Elastic Function

- Procedure to determine the parameters of elastic function
  1. Consult with regulatory agencies to choose a margin δ that represents a practically significant difference in treatment effect, e.g., δ = 0.2θ
  2. Simulate congruent case by simulating D with $\theta = \theta_0$ and obtain the corresponding T_0. Let T_{q0} denote the q_0th percentile of T_0.
  3. Simulate incongruent case by simulating D with $\theta = \theta_0 + \delta$ and obtain the corresponding T_1. Let T_{q1} denote the q_1th percentile of T_1.
  4. Solve (a, b) by setting

$$g(T_{q0}) = 1 \quad \text{(Full borrow)}$$

$$g(T_{q1}) = 0.01 \quad \text{(Discount } D_0 \text{ by a factor of 100)}$$

Page 27



An inconvenient fact

- **Theorem** For any method that borrows information from historical or other external data, dynamically or non-dynamically, the inflation of type I or II error is inevitable under finite samples, depending on whether historical or other external data under- or over-estimate the treatment effect of the control arm when compared to the current data.
- The reason is simple: even when the truth is that $\theta \neq \theta_0$, in finite samples, there is non-zero probability that the observed data D_0 and D are similar, thus triggers information borrowing and results inflated type I or II error

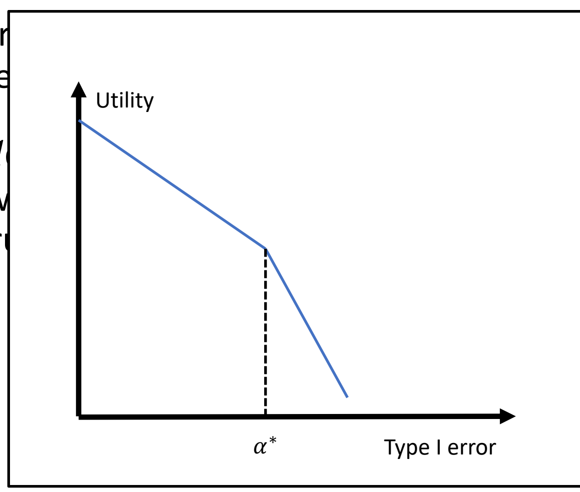
Wish gain in power, but no type I error inflation

Optimize the tradeoff

- Choose the bor timize power (under congrue ruence case)

$U(q$ $\alpha^*)$

where ρ is the pov e type I error under the incongr



Desirable Properties of Elastic Prior

- **Flexible:** readily accommodate the borrow/no borrow requirement prespecified by investigators or regulatory agency
- **Functional**: achieve adaptive information borrowing, i.e., full borrow (thus power gain) when D_0 and D are congruent, little borrow (thus little type I error inflation and bias) when D_0 and D are incongruent
- **Objective**: can be prespecified and included in the protocol before the trial starts
- **Realistic**: optimize the power-type-I-error tradeoff

Simulation

N=50 for historical, 25 for control, 50 for experimental;
Elastic prior: 0.5σ difference is regarded as practically incongruent

Mean			Type I error/Power			
Historical	Control	Experimental	**No Borrow**	**Commensurate prior**	**Power Prior**	**Elastic Prior**
1	1	1	5.0	5.0	5.0	5.0
1	1	1.5	66.3	91.6	88.3	93.6
0	1	1	5.0	14.6	30.0	7.3
2	1	1.5	66.3	57.6	37.8	72.6

Reference

- Jiang L, Nie L and Yuan Y (2020) Elastic priors to dynamically borrow information from historical data in clinical trials, arXiv:2009.06083

Contact: yyuan@mdanderson.org