Paper ET06

Leveraging Historical Control Data Through Adaptive Bayesian Hierarchical Models

Ajoy Mukhopadhyay, Novo Nordisk GBS, Bangalore, India
Debarghya Baul, Novo Nordisk GBS, Bangalore, India

Abstract

Introduction

Imagine being able to utilize what we've already learned from past studies—how could this transform the way we test new treatments? Bayesian methods enable seamless integration of historical data. Specifically, the use of synthetic controls allows sponsors to reduce or replace concurrent control arms, which can lower costs and speed up the drug development process. Additionally, with fewer controls required, these methods can reduce patient burden.

In Scope

- A comprehensive simulation study is utilized, assuming commensurate power priors for Gaussian response data, to illustrate the planning and execution of a trial.
- Model evaluation and sensitivity analysis are performed rigorously to ensure robustness in findings.
- Results comparison with traditional RCTs to discuss the benefit risk and regulatory aspects for this methodology.

Keywords: Innovative Bayesian Adaptive Design, Simulation based design, R packages

Disclaimer : Views expressed in the paper by the authors are the views, thoughts, and opinions belonging solely to the authors, and not necessarily to the authors' employer, organization or other group or individual.

Any data if used is from readily available package in CRAN or simulated. No data or confidential information belonging to Novo Nordisk has been used in this paper.

1. Introduction

Bayesian adaptive clinical trial methodology has evolved dramatically over the last few decades.

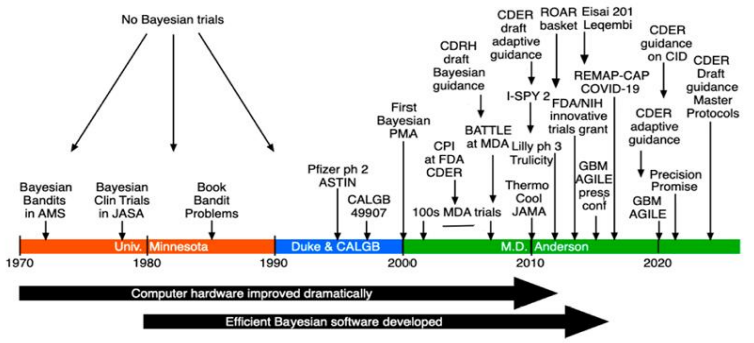


Figure : Timeline of Bayesian adaptive clinical trials. Donald Berry at the University of Minnesota, then Duke University, plus the Cancer & Leukemia Group B (CALGB), and then the University of Texas M.D. Anderson Cancer Center (MDACC) plus Berry Consultants. Arrows on clinical trials indicate time of the initial design. Numbers on the arrows indicate the relevant references.

Adaptive Bayesian Clinical Trials: The Past, Present, and Future of Clinical Research. DOI: 10.3390/jcm14155267

In context of borrowing from historical studies, the methods provide a flexible framework for incorporating external information into clinical trial analyses while formally accounting for uncertainty. In settings where relevant historical data are available, Bayesian borrowing can improve efficiency, reduce required sample sizes, and address ethical considerations related to unnecessary control-arm enrolment. However, inappropriate borrowing may lead to biased inference and inflation of Type-1 error when historical and current data are not sufficiently compatible Pocock et al. [1]

Several approaches have been proposed to enable principled borrowing from historical data while guarding against incompatibility. Commensurate prior model introduced by Hobbes et al. [2], explicitly link parameters from historical and current studies through a probabilistic structure, allowing the data to determine the degree of similarity. This framework formalizes the notation of exchangeability between studies and provides a coherent mechanism for adaptive borrowing driven by observed agreement between data sources.

In parallel, robust prior formulation, including robust meta-analytic-predictive (MAP) priors proposed by Schmidli et al.[3] , address prior-data conflict by mixing informative priors with weakly informative and vague components. These approaches emphasize protection against misspecification of historical information, ensuring that inference reverts to current-trial data when historical evidence proves unreliable.

In this paper, we propose a two-stage Bayesian adaptive design based on a robust mixture commensurate prior. The degree of borrowing is quantified through posterior inference and is used to guide pre-specified adaptations between trial stages.

To demonstrate the methodology, we consider a simulation study motivated by eGFR slope as a longitudinal endpoint in chronic kidney disease trials. The simulation evaluates the operating

characteristics of the design under varying degrees of historical-current compatibility, with particular emphasis on borrowing behaviour, control-arm sample-size implications and preservation of frequentist properties.

2. Statistical Methodology

2.1 Objectives and endpoints

We will consider conducting of a trial with an objective to demonstrate that the experimental drug delays the progression of renal impairment compared to placebo in subjects with chronic kidney disease (CKD).

Table 2.1: Objectives and endpoint

Objective	Endpoint	Time Frame
Progression of renal impairment	Annual rate of change in eGFR (mL/min/1.73m^2 per year) (CKD-EPI) (total eGFR slope)	From randomization (week 0) to 2 years with repeated measurements at every 26 weeks.

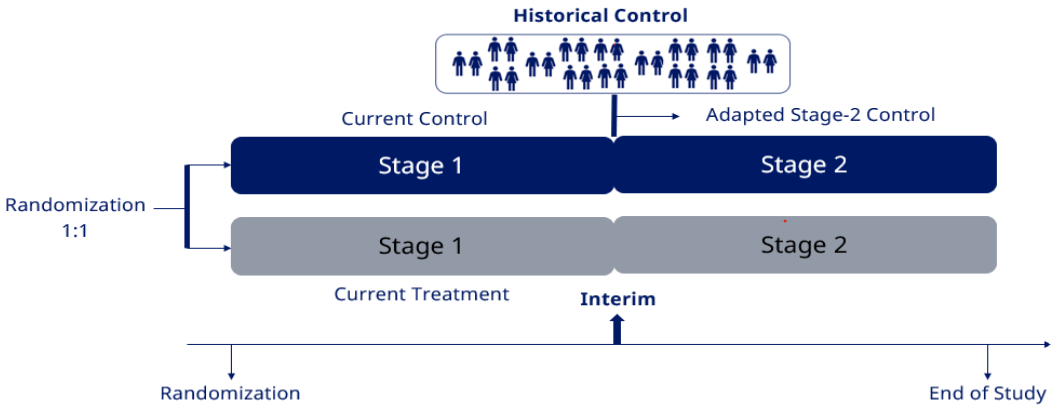
2.2 One Historical study

A previous CKD Phase 3 trial with eGFR slope as a surrogate endpoint is considered for the reference historical control.

2.3 Trial Design

The proposed design is a two-stage trial, in which stage-1 data are used to assess the degree of compatibility between historical and current control data. This assessment informs a pre-specified adaptation of the control-arm sample size in Stage-2. No adaptation of the control arm is considered if degree of compatibility between historical and current control data is significantly small.

Figure 2.3: Trial design diagram



2.4 Longitudinal Model for eGFR slope

A linear mixed-effects model with subject specific random intercept and random slope is considered for modelling eGFR for the historical control and current study.
Let's y_{ij} denotes the eGFR measurement for the subject i at visit j, observed at time t_{ij}

Current study (C)

$$y_{ij}^{C} = \beta_0^{C} + \beta_1^{C} t_{ij}^{C} + \gamma t_{ij}^{C} I(i \in Treatment) + b_{0i}^{C} + b_{1i}^{C} t_{ij}^{C} + \varepsilon_{ij}^{C}$$

Historical control (HC)

$$y_{ij}^{HC} = \beta_0^{HC} + \beta_1^{HC} t_{ij}^{HC} + b_{0i}^{HC} + b_{1i}^{HC} t_{ij}^{HC} + \varepsilon_{ij}^{HC}$$

Were:

- β_0^{C}, β_0^{HC} : fixed intercept
- β_1^{C}, β_1^{HC} : control slope
- γ : treatment difference in annual rate of change in eGFR (primary endpoint)
- b_{0i}^{C} , b_{0i}^{HC} : random intercept for the subject i
- b_{1i}^{C} , b_{1i}^{HC} : random slope for the subject i
- $\left(b_{0i}^{C}, b_{1i}^{C}\right)' \sim BVN(\tilde{0}, \Sigma_b^{C})$, $\left(b_{0i}^{HC}, b_{1i}^{HC}\right)' \sim BVN(\tilde{0}, \Sigma_b^{HC})$
- ε_{ij}^{C} , ε_{ij}^{HC}: independent residuals with variance σ_C^2 and σ_{HC}^2respectively

2.5 Commensurate Prior

The commensurate prior formulation of Hobbs et al. links the current and historical control parameters through a single commensurability parameter τ, allowing the data to modulate the degree of borrowing. While this approach adaptively adjusts borrowing strength, it relies on a continuous precision parameter and may still permit residual influence of historical information under substantial prior-data conflict.
To provide additional robustness, we extend the Hobbs commensurate prior by introducing a mixture formulation on the commensurability structure itself. Specifically, rather than assuming a single commensurate relationship between and, we allow the data to select between two commensurate regimes corresponding to strong and weak borrowing between historical control slope β_1^{HC} and current control slope β_1^{C}.

Formally, conditional on a latent indicator z, the current control slope is expressed as:

$$\beta_1^{C} \mid \beta_1^{HC}, z \sim \begin{cases} N\left(\beta_1^{HC}, \tau_{match}^{-1}\right); & z = 1 \\ N\left(\beta_1^{HC}, \tau_{no-match}^{-1}\right); & z = 0 \end{cases}$$

Where:

- $z = 1$: corresponds to the standard Hobbs commensurate prior with precision τ_{match}, representing strong commensurability.
- $z = 0$: represents a weakly commensurate prior with fixed low precision $\tau_{no-match}$, reflecting minimal borrowing.

The latent indicator follows $z \sim Bernoulli(\pi)$, with π representing the prior probability that the historical and current control slopes are strongly commensurate.

2.5 Posterior inference and Borrowing Assessment

Posterior inference jointly updates $(\beta_1^C,\ \beta_1^{HC},\ z\ , \tau)$ using the combined likelihood from historical and current data. The posterior distribution of the latent indicator *z* directly quantifies the evidence in favour of commensurability.

3. Simulation Study

3.1 Simulation Inputs

For simulation of historical and current study we followed the work by *wang and du [3]* on linear eGFR decline. For simplicity we have not considered any censoring due to drop-out or lost to follow-up within the trial duration.

3.1.1 Base Design

The base design to test the new treatment delays the progression of renal impairment compared to placebo, i.e.

$$H_0: \gamma = 0 \text{ against } H_1: \gamma > 0$$

We used a classical group-sequential design with one-sided α = 2.5% and a planned interim when 50% of the total participants have completed the treatment course.

Table 3.1.1: Base Design

Stage	Information Rate	α-Spending *Lan-DeMets (LD)*	Sample Size *Per-arm*
1	50%	0.001525	82
2	100%	0.025	164

***Slope difference under alternative: 2 ml/min/1.73m2 per year; slope standard deviation: 5.5; sample size is computed using RPACT package*

The target power is aimed at 90% and the slope difference under alternative is assumed to be 2 mL/min/m^2. The reference historical study is assumed to have 500 relevant control subjects.

3.1.2 Stage-2 Adaptation

Based on stage-1 data we will adapt the stage-2 control sample size as $N_{Ctr}^{Stage2} = 82 * (1 - w)$; where $w = median(z)$ represent the borrowing strength between historical and current control.

3.1.3 Prior Selection for the model parameters

We have considered non-informative priors on the model parameters to minimize the subjective beliefs and allowing the likelihood function (data) to dominate the posterior distribution. The commensurate prior for in-favour and not in-favour of borrowing is calibrated to detect small differences between historical and current control slope.

Table 3.1.3: Prior distribution

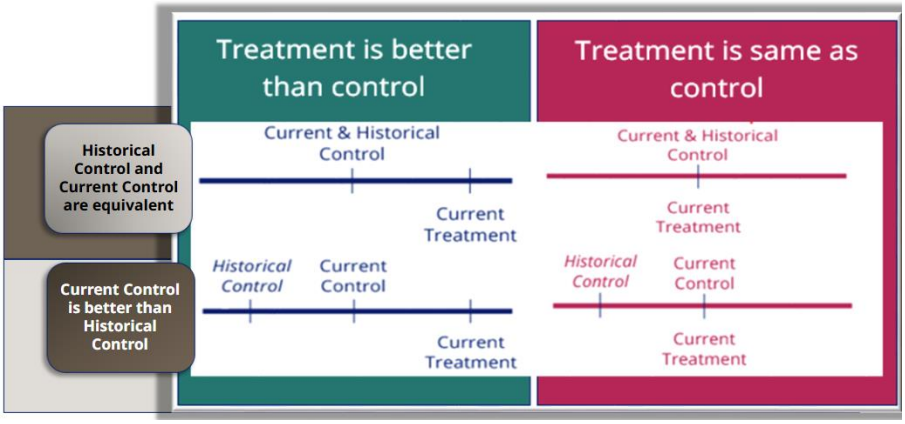
Prior Distributions		
Intercept and slope difference	$\beta_0^C,\ \beta_0^{HC}, \gamma$	$\mathcal{N}ormal(0, 20)$
Commensurate prior in favour of borrowing	τ_{match}	$Gamma(4, 0.08)$
Commensurate prior in favour of no-borrowing	$\tau_{no-match}$	$Gamma(0.1, 0.1)$
Prior probability in favour of borrowing	π	$Beta(1, 3)$
Random intercept variance	σ_{b0}^2	$StudentT(3, 0, 5)$
Random slope variance	σ_{b1}^2	$StudentT(3, 0, 2)$
Residual variance	$\sigma_C^2, \sigma_{HC}^2$	$StudentT(3, 0, 10)$

Note: No correlation between random effects

3.1.4 Sensitivity Analysis

We will consider the simulation under $H_0: \gamma = 0$ and $H_1: \gamma > 0$ with consideration of "*No Bias*" and "*Positive Bias*" between Current and Historical Control as described in the figure 3.1.2.

Figure 3.1.2: Simulation Scenarios



A parallel comparison with conventional group-sequential design will be performed assuming same stage-2 adapted control sample size to show the integrity between the two methods.

3.2 Simulation results

The general expectation is, the more the historical data is like the current control the less control samples we would need for stage-2 as it can be seen in the figure 3.2.1

Figure 3.2.1: Slope model estimates

Bayesian posterior estimates
Treatment is better than control – Current & Historical control are equivalent

Coefficient	Estimate (Median)	LL (2.5%)	UL (97.5%)	N (%)
Stage-1				
Intercept	52.63	50	55.53	
Control Slope	-2.34	-2.55	-2.1	
Borrowing strength (W)	0.87	0.01	1	
SD (Random Intercept)	17.01			
SD (Random Slope)	0.95			
SD (Residual)	1			
Stage-2 control sample size (% reduced from base)				11 (86.59%)
Stage-2				
Intercept	51.48	49.28	53.57	
Control Slope	-2.3	-2.49	-2.11	
SD (Random Intercept)	17.95			
SD (Random Slope)	1.01			
SD (Residual)	0.99			

Bayesian posterior estimates
Treatment is same as control – Current & Historical control are equivalent

Coefficient	Estimate (Median)	LL (2.5%)	UL (97.5%)	N (%)
Stage-1				
Intercept	51.3	48.63	54.08	
Control Slope	-2.35	-2.58	-2.09	
Borrowing strength (W)	0.89	0.01	1	
SD (Random Intercept)	17.64			
SD (Random Slope)	1.14			
SD (Residual)	1			
Stage-2 control sample size (% reduced from base)				9 (89.02%)
Stage-2				
Intercept	51.71	49.66	53.67	
Control Slope	-2.31	-2.51	-2.11	
SD (Random Intercept)	16.98			
SD (Random Slope)	1.09			
SD (Residual)	0.98			

Bayesian posterior estimates
Treatment is better than control – Current control is better than Historical control

Coefficient	Estimate (Median)	LL (2.5%)	UL (97.5%)	N (%)
Stage-1				
Intercept	52.96	50.26	55.38	
Control Slope	-2.33	-2.6	-2.06	
Borrowing strength (W)	0.01	0	1	
SD (Random Intercept)	17.03			
SD (Random Slope)	0.95			
SD (Residual)	1			
Stage-2 control sample size (% reduced from base)				81 (1.22%)
Stage-2				
Intercept	51.9	50.09	53.68	
Control Slope	-2.4	-2.58	-2.23	
SD (Random Intercept)	17.87			
SD (Random Slope)	1			
SD (Residual)	0.97			

Bayesian posterior estimates
Treatment is same as control – Current control is better than Historical control

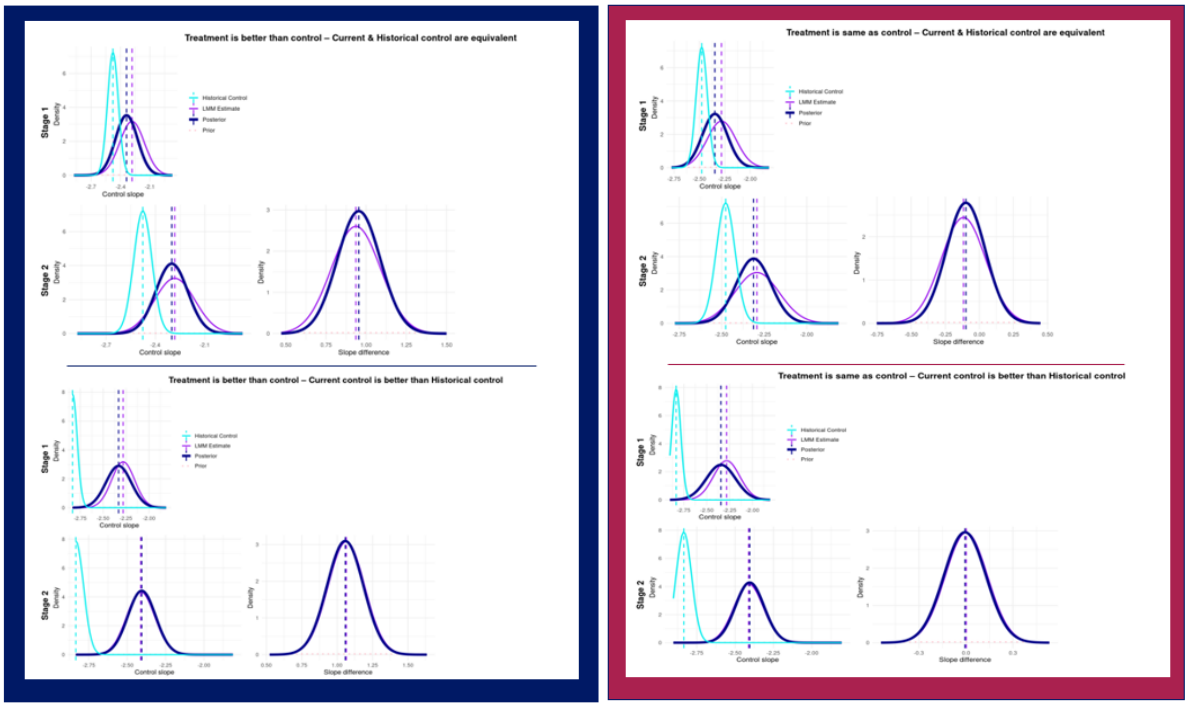
Coefficient	Estimate (Median)	LL (2.5%)	UL (97.5%)	N (%)
Stage-1				
Intercept	51.38	48.87	54.07	
Control Slope	-2.34	-2.66	-2.04	
Borrowing strength (W)	0.01	0	1	
SD (Random Intercept)	17.62			
SD (Random Slope)	1.14			
SD (Residual)	1			
Stage-2 control sample size (% reduced from base)				81 (1.22%)
Stage-2				
Intercept	51.72	49.68	53.56	
Control Slope	-2.41	-2.59	-2.22	
SD (Random Intercept)	17.09			
SD (Random Slope)	1.07			
SD (Residual)	0.97			

Results based on R simulated data

With the mixture distribution of τ we can see a minimum pull on the posterior distribution towards historical data figure 3.2.2 at stage-1. However, when the historical control differs from the current the posterior variance is comparatively higher than classical frequentist approach, and with a carefully chosen historical control this scenario is an indication of existence of an unknown confounder.

At stage-2 the adaptive sample size on control arm auto adjusts from (low) ➔ (earlier planned group-sequential sample size) as the likelihood of similarity between observed data and historical data drops. The final analysis is based on combined stage-1, stage-2 data and the posterior estimate of τ from stage-1.

Figure 3.2.2: Distribution — Prior, Posterior & Data

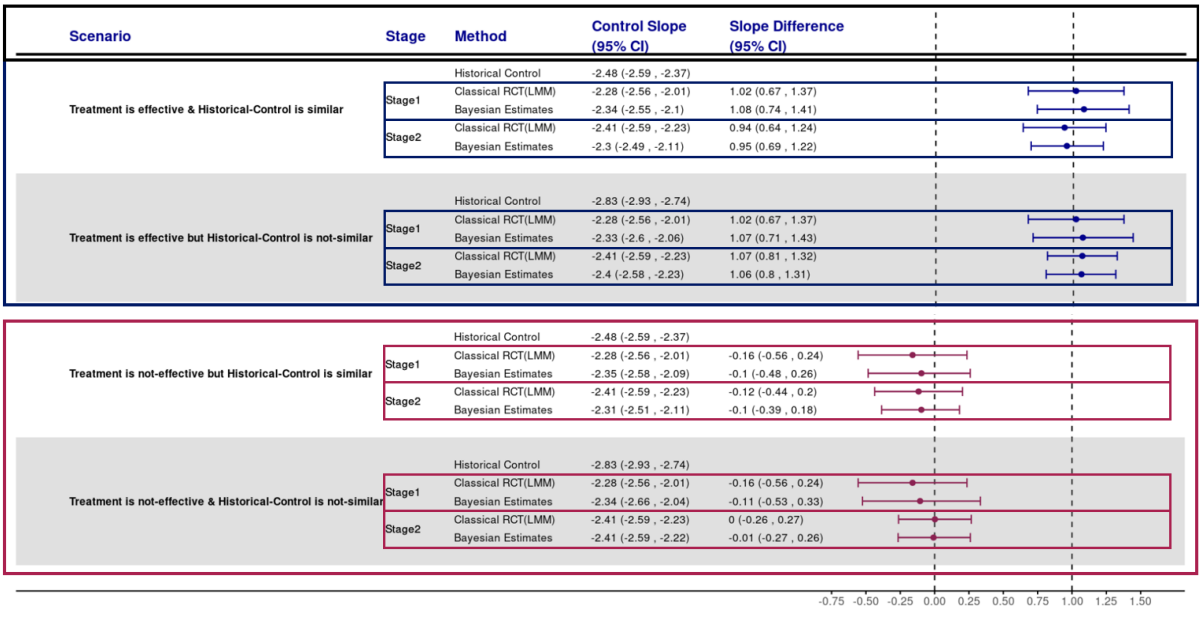


***Posterior density curve: Normal density with mean as the median of the posterior samples*

Results based on R simulated data

Figure 3.2.3: Treatment effect – difference annual rate of change in eGFR (mL/min/1.73m^2 per year)

Scenario	Stage	Method	Control Slope (95% CI)	Slope Difference (95% CI)
Treatment is effective & Historical-Control is similar		Historical Control	-2.48 (-2.59 , -2.37)	
	Stage1	Classical RCT(LMM)	-2.28 (-2.56 , -2.01)	1.02 (0.67 , 1.37)
		Bayesian Estimates	-2.34 (-2.55 , -2.1)	1.08 (0.74 , 1.41)
	Stage2	Classical RCT(LMM)	-2.41 (-2.59 , -2.23)	0.94 (0.64 , 1.24)
		Bayesian Estimates	-2.3 (-2.49 , -2.11)	0.95 (0.69 , 1.22)
Treatment is effective but Historical-Control is not-similar		Historical Control	-2.83 (-2.93 , -2.74)	
	Stage1	Classical RCT(LMM)	-2.28 (-2.56 , -2.01)	1.02 (0.67 , 1.37)
		Bayesian Estimates	-2.33 (-2.6 , -2.06)	1.07 (0.71 , 1.43)
	Stage2	Classical RCT(LMM)	-2.41 (-2.59 , -2.23)	1.07 (0.81 , 1.32)
		Bayesian Estimates	-2.4 (-2.58 , -2.23)	1.06 (0.8 , 1.31)
Treatment is not-effective but Historical-Control is similar		Historical Control	-2.48 (-2.59 , -2.37)	
	Stage1	Classical RCT(LMM)	-2.28 (-2.56 , -2.01)	-0.16 (-0.56 , 0.24)
		Bayesian Estimates	-2.35 (-2.58 , -2.09)	-0.1 (-0.48 , 0.26)
	Stage2	Classical RCT(LMM)	-2.41 (-2.59 , -2.23)	-0.12 (-0.44 , 0.2)
		Bayesian Estimates	-2.31 (-2.51 , -2.11)	-0.1 (-0.39 , 0.18)
Treatment is not-effective & Historical-Control is not-similar		Historical Control	-2.83 (-2.93 , -2.74)	
	Stage1	Classical RCT(LMM)	-2.28 (-2.56 , -2.01)	-0.16 (-0.56 , 0.24)
		Bayesian Estimates	-2.34 (-2.66 , -2.04)	-0.11 (-0.53 , 0.33)
	Stage2	Classical RCT(LMM)	-2.41 (-2.59 , -2.23)	0 (-0.26 , 0.27)
		Bayesian Estimates	-2.41 (-2.59 , -2.22)	-0.01 (-0.27 , 0.26)



Results based on R simulated data

4. Conclusions

In this work we showed the borrowing from one historical study however without loss of generality the methodology can be extended for multiple historical studies, alternatively one can take subset of participants from the study or studies by propensity score matching to get further refinement in the historical data on the same line as the methodology proposed by Jiang et al [5]

Another possible improvement can be done at the planning stage sample size computation by incorporating meta-analytic prior and effective sample size instead of using group-sequential sample size, Schmidli et al [3].

In favourable situations when historical control and current control are similar, the posterior estimate of treatment effect can provide much accurate estimates at stage-1 itself which can help to escalate later phases of drug development.

While the methodology shows promising results by providing data driven flexibility of incorporating previous knowledge into trial planning, some precautions need to take into consideration on historical study or prior selection. The primary focus of this paper is to show borrowing methodology side by side in comparison to a classical RCT trial without stating any trial success criteria for efficacy, to decide the efficacy threshold rigorous simulation studies need to be performed to ensure type-1 error is maintained, FDA guidelines on Bayesian methodology in Clinical Trials, 2026 [6].

5. Appendix

Stan code for Stage-1 analysis

```
stage1_stan <- "
data {
    int<lower=1> N_hist;
    int<lower=1> N_subj_hist;
    int<lower=1> id_hist[N_hist];
    vector[N_hist] time_hist;
    vector[N_hist] eGFR_hist;

    int<lower=1> N_current;
    int<lower=1> N_subj_current;
    int<lower=1> id_current[N_current];
    int<lower=0,upper=1> treatment[N_current];
    vector[N_current] time_current;
    vector[N_current] eGFR_current;
}
```

← Input data block :
Stage-1 data, Historical control

```
parameters {
// LMM model parameters Historical
real beta0_0;
real beta1_0;
vector[N_subj_hist] u0_hist;
vector[N_subj_hist] u1_hist;

// LMM model parameters Current trial
real beta0;  // intercept
real beta1;  // control slope - commensurate
real gamma;  // treatment effect on slope
vector[N_subj_current] u0_current;
vector[N_subj_current] u1_current;


// Commensurability parameters (Mixture prior for tau)
real<lower=0> tau_match;
real<lower=0> tau_nomatch;
real<lower=0, upper=1> pi_mix;
//real<lower=0> tau_slope;

// variance components
real<lower=0> sigma_hist;
real<lower=0> sigma_current;
real<lower=0> sigma_u0_hist;
real<lower=0> sigma_u1_hist;
real<lower=0> sigma_u0_current;
real<lower=0> sigma_u1_current;
}
```

← Model parameters block

```
model {
  // Priors for fixed effects
  // historical control
  beta0_0 ~ normal(0, 20);
  beta1_0 ~ normal(0, 20);

  // current
  beta0 ~ normal(0, 20);
  gamma ~ normal(0, 20);

  // variance prior
  sigma_current ~ student_t(3, 0, 10);
  sigma_hist ~ student_t(3, 0, 10);
  sigma_u0_hist ~ student_t(3, 0, 5);
  sigma_u1_hist ~ student_t(3, 0, 2);
  sigma_u0_current ~ student_t(3, 0, 5);
  sigma_u1_current ~ student_t(3, 0, 2);
```

← Model block

```
  // random effect
  u0_hist ~ normal(0, sigma_u0_hist);
  u1_hist ~ normal(0, sigma_u1_hist);
  u0_current ~ normal(0, sigma_u0_current);
  u1_current ~ normal(0, sigma_u1_current);

  // Mixture Commensurate priors
  pi_mix ~ beta(1, 3);
  tau_match ~ gamma(4, 0.08);   // strong borrowing
  tau_nomatch ~ gamma(0.1, 0.1); // weak borrowing
  target += log_mix(pi_mix,
                    normal_lpdf(beta1|beta1_0, inv_sqrt(tau_match)),
                    normal_lpdf(beta1|beta1_0, inv_sqrt(tau_nomatch)));


  // Likelihood: Current trial
  for (n in 1:N_current) {
    real mu_current = beta0
      + beta1 * time_current[n]
      + gamma * time_current[n] * treatment[n]
      + u0_current[id_current[n]]
      + u1_current[id_current[n]] * time_current[n];
    eGFR_current[n] ~ normal(mu_current, sigma_current);
  }


  // Likelihood: Historical data
  for (n in 1:N_hist) {
    real mu_hist = beta0_0
      + beta1_0 * time_hist[n]
      + u0_hist[id_hist[n]]
      + u1_hist[id_hist[n]] * time_hist[n];
    eGFR_hist[n] ~ normal(mu_hist, sigma_hist);
  }
}
```

← Model block

```
generated quantities {
  // posterior borrowing probability
  real lp_match = normal_lpdf(beta1|beta1_0, inv_sqrt(tau_match));
  real lp_nomatch = normal_lpdf(beta1|beta1_0, inv_sqrt(tau_nomatch));
  real z = 1/(1+((1-pi_mix)*exp(lp_nomatch))/(pi_mix*exp(lp_match)));
  real borrowing_strength = z;
  real tau = z*tau_match + (1-z)*tau_nomatch;
}
"
```

← Adaptive weight

6. References

[1] Pocock SJ. The combination of randomized and historical controls in clinical trials. Journal of chronic Disease, 1976; 29: 175-188 [PubMed: 770493]

[2] Hobbs BP, Carlin BP, Mandrekar SJ, Sargent DJ. Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. Biometrics. 2011 Sep;67(3):1047-56. doi: 10.1111/j.1541-0420.2011.01564.x. Epub 2011 Mar 1. PMID: 21361892; PMCID: PMC3134568.

[3] Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B. Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics. 2014 Dec;70(4):1023-32. doi: 10.1111/biom.12242. Epub 2014 Oct 29. PMID: 25355546.*

[4] Estimand framework development for eGFR slope estimation and comparative analyses across various estimation methods, Tuo Wang and Yu Du, Global Statistical Sciences, Eli Lilly and Company.

[5] Jiang L, Thall PF, Yan F, Kopetz S, Yuan Y. BASIC: A Bayesian adaptive synthetic-control design for phase II clinical trials. Clin Trials. 2023 Oct;20(5):486-496. doi: 10.1177/17407745231176445. Epub 2023 Jun 14. PMID: 37313712; PMCID: PMC10504821.

[6] Draft guidance for use of Bayesian methodology in Clinical Trials of Drug and Biological Products Guidance for Industry, Division of Drug Information Centre for Drug Evaluation and Research Food and Drug Administration(FDA)

7. **CONTACT INFORMATION**

Your comments and questions are valued and encouraged. Contact the authors at:

First Author Name: Ajoy Mukhopadhyay
Company: Novo Nordisk GBS
Address: 560066, Bangalore, India
Work Phone: 7795841349
Email : ojmp@novonordisk.com

Second Author Name: Debarghya Baul
Company: Novo Nordisk GBS
Address: 560066, Bangalore, India
Work Phone: 8105869142
Email : dghb@novonordisk.com