Bayesian methods for placebo borrowing in master protocols

Will Landau, Albert Man, Phebe Kemmer, Saptarshi Chatterjee, Faith Bian, Heather Zhao, and Zhangchen Zhao.

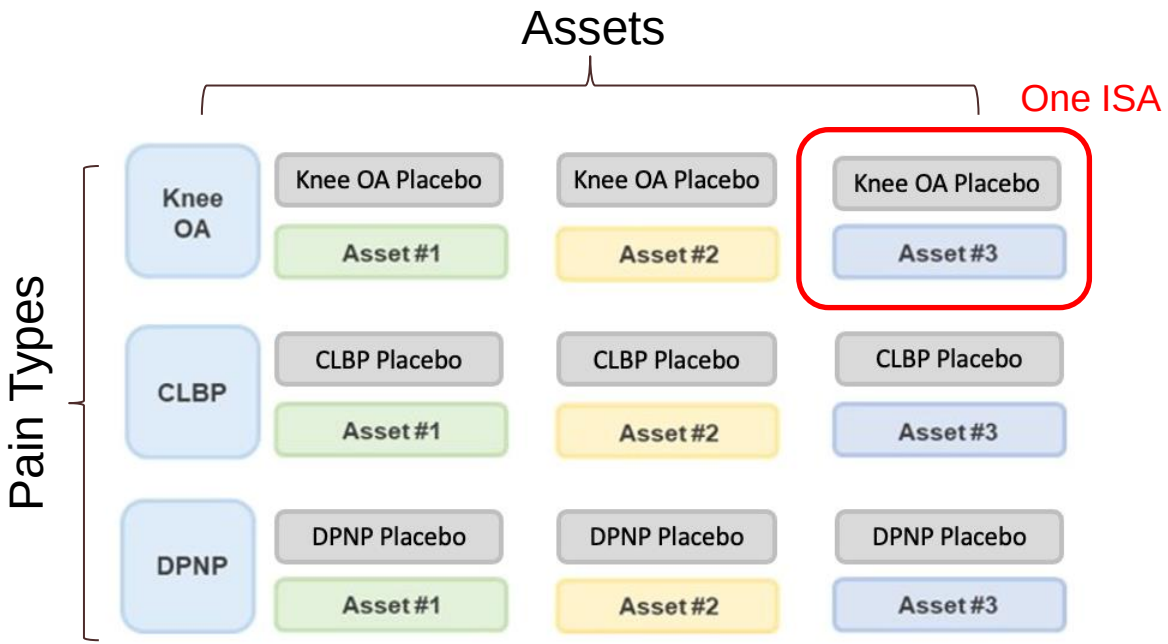
Lilly

Outline

- Chronic Pain Master Protocol
- Placebo borrowing methods
- A longitudinal hierarchical model for placebo borrowing
- Implementation challenges and solutions

Chronic Pain Master Protocol (CPMP)

Chronic pain master protocol (CPMP) Overview



- Each of the 9 cells in this grid represents an intervention-specific appendix, or ISA, of the master protocol.
- **Each ISA is its own clinical trial**: a parallel randomized controlled phase 2a (proof of concept) design with 1 active arm and its own placebo arm.
- The primary analysis is the change from baseline in average pain intensity (API) measured via the numeric rating scale (NRS), collected daily using an eDiary device.

Placebo borrowing

Placebo borrowing

- Leverage placebo data from historical ISAs to analyze the current ISA.
- Potential to improve operating characteristics such as power.
- Operational conditions that allow placebo borrowing in CPMP:
  - Multiple candidate drugs studied within each pain type.
  - Statistical exchangeability.
    - Standardized data collection, pain scales, randomization, visit structure.
    - Same investigative sites across ISAs*.

*ISA = intervention-specific appendix, a clinical trial within the master protocol.

Caveat

- What happens when historical placebo data is not commensurate with current data?
- Which better reflects the truth about the drug, the current data or the historical data?
  - Can depend on operational issues (e.g. different modes of administration), or the responses in the current vs historical datasets could systematically differ for unknown reasons.
  - To minimize human subjectivity, it is critical to plan for these scenarios before accessing the data.
- Dynamic borrowing methods automatically decide how strongly to borrow.
  - Often done by comparing historical outcomes to current outcomes.
  - Also possible to use baseline covariates.

Historical borrowing methods

- **Pooling**: append the historical data onto the current data for analysis.
- **Test-then-pool**: conduct a hypothesis test to compare the historical and current data. Pool if the null hypothesis is not rejected.
- **Propensity scores**:
  - Use baseline covariates to select historical control patient data similar to the current control data.
  - Can be included in the current control arm (data augmentation) or combined with another technique like the MAP prior.
- **Power priors:**
  - **Conditional power priors**: historical data informs a Bayesian prior on the control group, and part of the prior density is raised to a fixed power in order to mitigate the influence of the historical data.
  - **Modified power priors**: put a prior on the power parameter above.
  - **Commensurate power priors**: adjust the prior on the power parameter based on the strength of evidence that the historical data is commensurate with the current data.
- **Meta-analytic predictive (MAP) priors**: construct a prior distribution for the current ISA using the predictive distribution of the current ISA from a meta-analysis hierarchical model on the historical studies.
- **Mixture modeling**: assign a mixture prior with each mixture component corresponding to a historical ISA. Intuitive and interpretable for 1 historical ISA.
- **Hierarchical modeling**: borrow information across studies to shrink placebo group means towards a common grand mean. Intuitive and interpretable for multiple historical ISAs.

Longitudinal hierarchical borrowing

Hierarchical model

$$y_k \sim \text{MVN}((X_\alpha)_k \cdot \alpha + (X_\delta)_k \cdot \delta + (X_\beta)_k \cdot \beta,\ I_{N_k} \otimes \Sigma_k)$$

$$\alpha_{kt} \overset{\text{ind}}{\sim} \text{Normal}(\mu_t, \tau_t^2)$$

$$\mu_t \overset{\text{ind}}{\sim} \text{Normal}(0, s_\mu^2)$$

$$\tau_t \overset{\text{ind}}{\sim} \text{Uniform}(0, s_\tau)$$

$\ldots$

- y_k: vector of observed patient-level outcomes of ISA* k.
- α: vector of α_{kt} parameters.
- α_{kt}: placebo mean response of ISA k time t.
- $(X_\alpha)_k$: indicator matrix to match the correct α_{kt} to each patient and time point.
- μ_t grand mean placebo response at time t.
- τ_t: standard deviation of the placebo means, controls the strength of dynamic borrowing.

*ISA = intervention-specific appendix, a clinical trial within the master protocol.

Hierarchical model

$$y_k \sim \text{MVN}((X_\alpha)_k \cdot \alpha + (X_\delta)_k \cdot \delta + (X_\beta)_k \cdot \beta,\ I_{N_k} \otimes \Sigma_k)$$

$$\dots$$

$$\delta_{dt} \overset{\text{ind}}{\sim} \text{Normal}(0, s_\delta^2)$$

$$\beta_b \overset{\text{ind}}{\sim} \text{Normal}(0, s_\beta^2)$$

- y_k: vector of observed patient-level outcomes of ISA k.
- δ: vector of δ_{dt} parameters.
- δ_{dt}: group mean response of active arm d time t.
- $(X_\delta)_k$: indicator matrix to match the correct δ_{kt} to each patient and time point.
- β: vector of β_b parameters.
- β_b: baseline covariate fixed effect parameter.
- $(X_\beta)_k$: model matrix of baseline covariates.

Hierarchical model

$$y_k \sim \text{MVN}((X_\alpha)_k \cdot \alpha + (X_\delta)_k \cdot \delta + (X_\beta)_k \cdot \beta,\ I_{N_k} \otimes \Sigma_k)$$

$$\dots$$

$$\Sigma_k = (I_T \sigma_k)\, \Lambda_k \Lambda_k' \,(I_T \sigma_k)$$

$$\sigma_{k1}, \dots, \sigma_{kT} \overset{\text{ind}}{\sim} \text{Uniform}(0, s_\sigma)$$

$$\Lambda_k \Lambda_k' \sim \begin{cases} \text{LKJ}(\text{shape} = s_\lambda,\ \text{order} = T) & m_k = 1 \\ \text{AR}(1)(T, \rho_k) & m_k = 2 \\ I_T & m_k = 3 \end{cases}$$

$$\rho_k \overset{\text{ind}}{\sim} \text{Uniform}(-1, 1) \qquad (\text{only for } m_k = 2)$$

- y_k: vector of observed patient-level outcomes of ISA k.
- N_k: number of patients in ISA k.
- I_{Nk}: identity matrix with N_k rows.
- $\otimes$: Kronecker product.
- Σ_k: longitudinal covariance matrix block of ISA k. Has covariances among time points within patients.
- Λ_k: Cholesky factor of the longitudinal correlation matrix block of ISA k.
- σ_k: vector of σ_{kt} parameters.
- σ_{kt}: residual standard deviation of ISA k time t.
- ρ_k: correlation between adjacent times within patients for ISA k (AR(1) only).
- m_k: choice of covariance structure of ISA k: 1 for unstructured/LKJ, 2 for AR(1), and 3 for diagonal.

Benchmark models to quantify borrowing

Model	Borrowing strength	Model of placebo response	Description
Hierarchical	Dynamic borrowing	$\alpha_{kt} \overset{\text{ind}}{\sim} \text{Normal}\left(\mu_t, \tau_t^2\right)$ $\mu_t \overset{\text{ind}}{\sim} \text{Normal}\left(0, s_\mu^2\right)$ $\tau_t \overset{\text{ind}}{\sim} \text{Uniform}\left(0, s_\tau\right)$	At each time t, the placebo means α_{kt} of each ISA k share a common mean parameter μ_t and common standard deviation τ_t.
Independent	No borrowing	$\alpha_{kt} \overset{\text{ind}}{\sim} \text{Normal}\left(0, s_\alpha^2\right)$	Placebo means α_{kt} are independent.
Pooled	Full borrowing	$\alpha_t \overset{\text{ind}}{\sim} \text{Normal}\left(0, s_\alpha^2\right)$	All studies k share a common placebo mean α_t at time t.

Quantify borrowing

- Want to empirically measure the strength of borrowing of the hierarchical model.
- One approach is to compare the results of the hierarchical model against two benchmark models:
  - Independent: like the hierarchical model, but with independent diffuse priors on the placebo means.
  - Pooled: like the independent model, but with a single shared placebo mean.
- Notation:
  - $u_t^{(m)}$ posterior mean of the placebo mean.
  - $v_t^{(m)}$: posterior variance of the placebo mean.
  - $n_t^{(m)}$: number of observed data points.
  - t: time point.
  - m: 1 for the hierarchical model, 2 for the independent model, and 3 for the pooled model.

Mean shift ratio

$$\frac{u_t^{(1)} - u_t^{(2)}}{u_t^{(3)} - u_t^{(2)}}$$

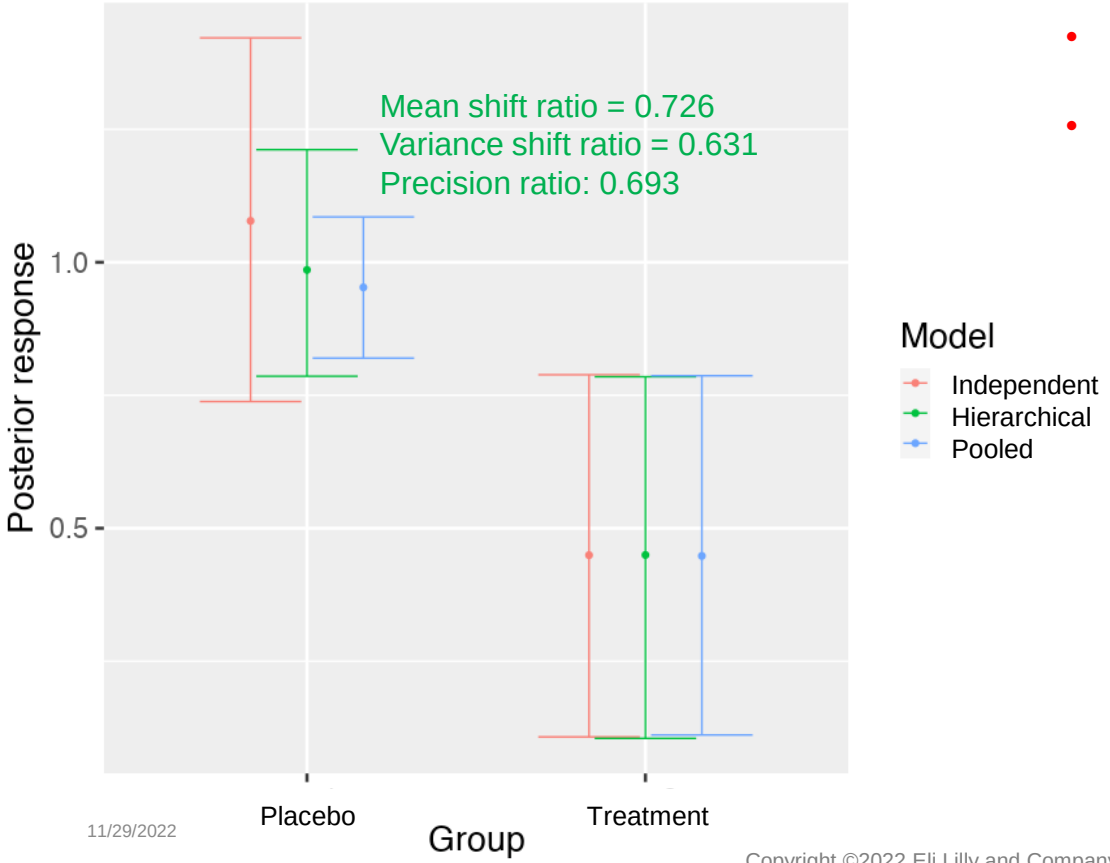
Variance shift ratio

$$\frac{v_t^{(1)} - v_t^{(2)}}{v_t^{(3)} - v_t^{(2)}}$$

Precision ratio

$$\frac{\frac{1}{\tau_t^2}}{\frac{1}{\tau_t^2} + \frac{n_t}{\sigma_t^2}}$$

Example results for a single time point



- Placebo borrowing is about 63%-73% according to the borrowing metrics.
- The mean shift ratio and variance shift ratio are visible in the graph.

Implementation

Software

- Published R package {historicalborrowlong} on CRAN.*
  - Hierarchical, independent, and pooled models.
- The package has Stan code included using R packages {rstan} and {rstantools}.
- We could not use JAGS or NIMBLE for longitudinal modeling.
  - The data model for each patient is multivariate normal, and neither JAGS nor NIMBLE supports partially missing data in this situation.
- Because of the custom hierarchical priors and unstructured covariance matrices, we could not use existing high-level packages like {MCMCglmm} or {rstanarm} (although the latest version of {brms} would have been an option).

*The {historicalborrow} package implements non-longitudinal versions of these models, as well as a mixture model. Also on CRAN.

Main challenges and solutions

Challenges

1. When baseline covariates are included straightforwardly, the hierarchical model only borrows a conditional subset of patients, which incorrectly weakens the strength of borrowing.

2. The initial implementation of the hierarchical model was slow to accrue effective samples (several hours of runtime to satisfy convergence diagnostics).

3. For some analyses, the mean shift ratio is not always between 0 and 1.

Solutions

1. Within each ISA*, center the columns of the model matrix for baseline covariate fixed effects. This ensures the placebo mean parameters are not conditional on a subset of the data.

2. Use a non-centered parameterization and vectorize the data model. Achieved a 22x speedup (23.5 minutes of runtime) from switching to the non-centered parameterization alone.

3. This behavior seems linked to elaborate covariance structures, informative priors, and cases when the independent and pooled models agree. Work on this issue is still ongoing.

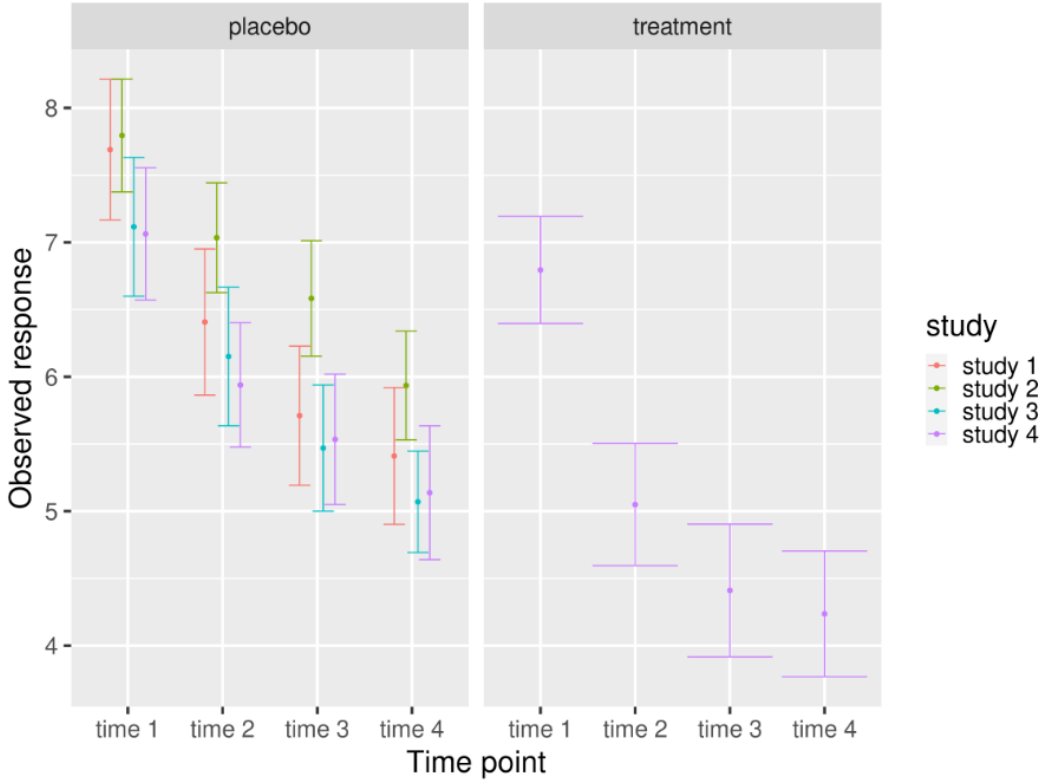
*ISA = intervention-specific appendix, a clinical trial within the master protocol.

Consequences of uncentered baseline covariates on borrowing

- Consider a **non-centered** baseline covariate:
  - 2 investigative sites: one in Indiana, another in New Jersey.
  - $(X_\beta)_k = (1, 0, 0, 1, 0, 1, ...)^T$: 0 for Indiana patients, 1 for New Jersey patients.
- Example observations:
  - y_{121} (ISA 1 patient 2 time 1) is from an Indiana patient: $\longrightarrow \mathrm{E}(y_{121}) = \alpha_{11}$
  - y_{131} (ISA 1 patient 3 time 1) is from a New Jersey patient: $\longrightarrow \mathrm{E}(y_{131}) = \alpha_{11} + \beta_1$
- This lack of centering **impairs borrowing**.
  - α_{11} (ISA 1 time 1) becomes the placebo mean ***at Indiana***, not the placebo mean overall.
  - Borrowing only occurs for Indiana placebo patients, not for all placebo patients.
- **Solution: subtract each column by its mean**.
  - $(X_\beta)_k = (0.46, -0.54, -0.54, 0.46, -0.54, 0.46, ...)^T$ (assuming 54% of patients are from New Jersey).
  - Expected values shift:
    $$\mathrm{E}(y_{121}) = \alpha_{11} - 0.54 \cdot \beta_1$$
    $$\mathrm{E}(y_{131}) = \alpha_{11} + 0.46 \cdot \beta_1$$
  - The mean of $(X_\beta)_k$ becomes 0.
  - The regression reference level rightly becomes the placebo mean of the ISA.

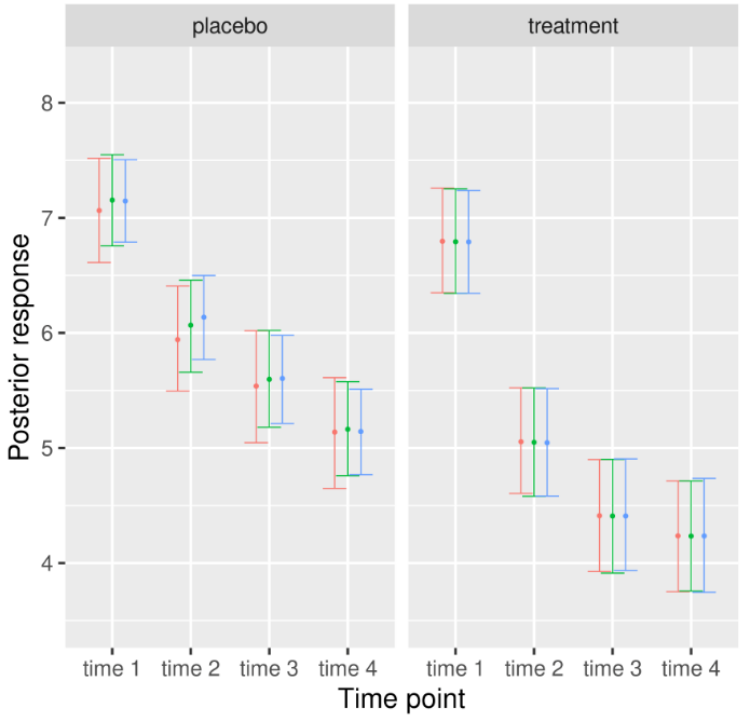
Example: simulated data

- Demonstration on an example dataset with 4 studies in chronological order.
  - Historical ISA: studies 1, 2, 3
  - Current ISA: study 4
- 100 virtual patients per ISA per arm randomized evenly over 4 investigative sites (the only baseline covariate).
- Simulated from the independent model with fixed placebo means α, fixed treatment means δ, and AR(1) covariances (75% correlation).
- The simulated response variable is pain on a numeric rating scale, and a lower score means less pain.

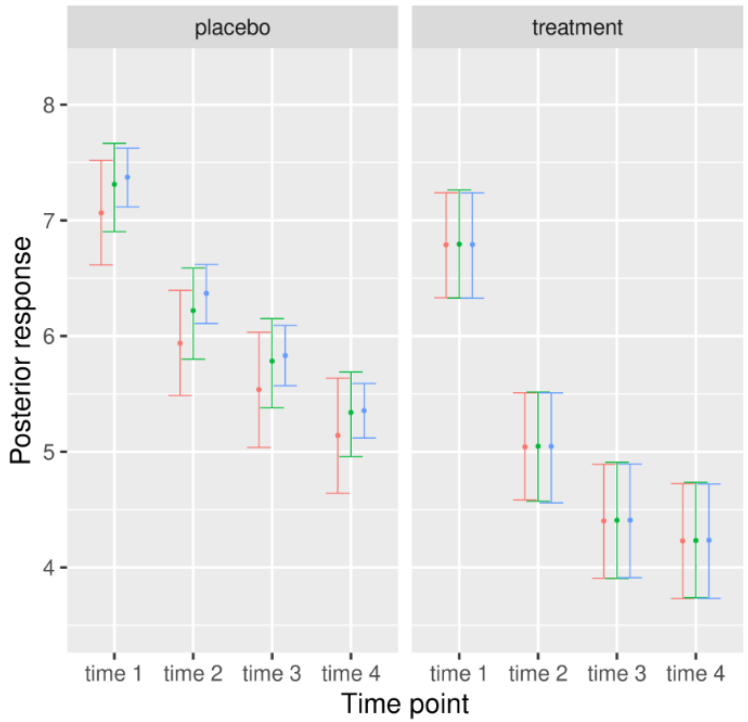


Covariate centering within each ISA

Without centering ✗



With centering ✓



- Failing to center within ISA incorrectly weakens borrowing because α is conditional on a specific reference level.
- Centering globally (not within ISA) produces similar problems/results as not centering at all.

Successful computational techniques for the hierarchical model in Stan

- Non-centered parameterization for placebo mean parameters: https://mc-stan.org/docs/stan-users-guide/reparameterization.html.
  - Transforms the scale of the density to be less variable across different regions of the sample space.
  - Allows Hamiltonian Monte Carlo (HMC) tuning parameters such as leapfrog step size to be better adjusted to more regions of the density.
  - 22x speedup for 4 chains and 8000 iterations per chain (4000) warmup (8.6 hours => 23.5 minutes).
  - In the code: $\alpha_{kt} = \alpha^*_{kt} \cdot \tau_t + \mu_t;\ \alpha^*_{kt} \overset{\text{ind}}{\sim} \text{Normal}(0, 1)$
- Vectorize the data model.
  - y ~ multi_normal_cholesky(vectors mu, matrix L)
  - mu is an array of vectors, where each vector is the vector of longitudinal data of a patient.
- Matrices
  - Instead of multiplying by sparse binary matrices $(X_\alpha)_k$ and $(X_\delta)_k$, use index vectors on α and δ.
  - Multiply individual blocks $(X_\beta)_k$ instead of the full baseline covariates model matrix X_β.
    - Prior to multiplying by β, select only the baseline (time = 1) rows of $(X_\beta)_k$ (these are baseline covariates, not time-varying).

Summary

Conclusions

- Summary
  - Placebo borrowing has the potential to increase power and decrease sample size in master protocols.
  - We implemented dynamic placebo borrowing for a master protocol in chronic pain and used our models in the topline analyses of several ISAs.
    - Also implemented non-longitudinal hierarchical and mixture models (not discussed here).
  - We propose simple metrics to quantify the strength of historical borrowing.
  - In the hierarchical and pooled borrowing models discussed here, it is important to center baseline covariates so borrowing is not conditional on an arbitrary subset of patients.
  - We chose Stan to implement the models and developed an efficient R package to carry out the standardized analyses of our chronic pain trials.
- Future work
  - Resolve practical edge cases applying borrowing metrics like the mean shift ratio.
  - Alternative placebo borrowing methods:
    - Bayesian model averaging of the independent and pooled models.
    - Tipping point analysis.
    - Longitudinal mixture model.

References

- Hobbs BP, Sargent DJ, Carlin BP. Commensurate Priors for Incorporating Historical Information in Clinical Trials Using General and Generalized Linear Models. Bayesian Anal. 2012 Aug 28;7(3):639-674. doi: 10.1214/12-BA722. PMID: 24795786; PMCID: PMC4007051.
- Hobbs BP, Carlin BP, Mandekar SJ, Sargent DJ. Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. Biometrics. 2011; 67:1047– 1056.
- Ibrahim JG, Chen MH. Power prior distributions for regression models. Statistical Science 2000; 15:46–60.
- Li, Jianghao and Du, Yu and Liu, Huayu and Yi, Yanyao. A matching design for augmenting a randomized clinical trial with external control. arXiv 2022, https://arxiv.org/abs/2203.10128.
- Lin J, Gamalo-Siebers M, Tiwari R. Propensity score matched augmented controls in randomized clinical trials: A case study. Pharm Stat. 2018 Sep;17(5):629-647. doi: 10.1002/pst.1879. Epub 2018 Jul 31. PMID: 30066459.
- Liu, M, Bunn, V, Hupf, B, Lin, J, Lin, J. Propensity-score-based meta-analytic predictive prior for incorporating real-world and historical data. *Statistics in Medicine*. 2021; 40: 4794– 4808. https://doi.org/10.1002/sim.9095
- Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter D (2010). “Summarizing Historical Information on Controls in Clinical Trials.” Clinical Trials, 7(1), 5–18.
- Pocock SJ. The combination of randomized and historical controls in clinical trials. Journal of Chronic Diseases 1976; 29:175–188.
- 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, 70: 1023-1032. https://doi.org/10.1111/biom.12242
- Spiegelhalter D, Abrams K, Myles J (2004). Bayesian Approaches to Clinical Trials and Health-Care Evaluation. John Wiley & Sons.
- Stan Development Team (2022). Stan Reference Manual. Version 2.30. https://mc-stan.org/docs/reference-manual/index.html
- Thompson L, Chu Jianxiong, Xu J et al. (2021) Dynamic borrowing from a single prior data source using the conditional power prior, Journal of Biopharmaceutical Statistics, 31:4, 403-424, DOI: 10.1080/10543406.2021.1895190
- Viele K, Berry S, Neuenschwander B et al. Use of historical control data for assessing treatment effects in clinical trials. Pharm Stat. 2014 Jan-Feb;13(1):41-54. doi: 10.1002/pst.1589. Epub 2013 Aug 5. PMID: 23913901; PMCID: PMC3951812.