MUCE: A Bayesian Hierarchical Model for Multiple Cohort Expansion

Jiaying Lyu[1,2], Wentian Guo[2], Yuan Ji[1,2,3,*] Email: yuanji@laiyaconsulting.com
[1]Fudan University; [2]Laiya Consulting Inc.; [3]The University of Chicago

Introduction

Background In 2018, FDA released a draft guidance for FIH multiple cohort expansion trial[1], which recommends cohort expansion on multiple disease-specific sub-populations/indications even on multiple doses simultaneously to assess the anti-tumor activities on multiple indications and doses after the dose-escalation phase.

Considered that treatment effects in various indications may provide some information about treatment effects in other indications, some Bayesian hierarchical models (BHM) have been applied in Phase II oncology trials to improve the statistical power.

Purpose We proposed a novel Bayesian hierarchical model to evaluate the treatment effect for multiple cohort expansion trials, called MUCE, and compared it with other three designs: Simon's 2-stage design, Berry's BHM[2] and Chu and Yuan's CBHM[3].

Bayesian Hierarchical Models

Consider a trial that evaluates the efficacy of a new drug in J cohorts. Let q_j, n_j and y_j denote the response rate, the number of patients and responders at cohort j, respectively. The objective of the trial to test whether the targeted drug is effective in each cohort,

$$H_0 : q_j \leq q_0 \quad \text{versus} \quad q_j \geq q_1$$

where q_0 is the reference rate , and q_1 is the target response rate.

BHM Berry et. al [2] assumed y_j follows a hierarchical model,

$$\begin{aligned} y_j \mid n_j, q_j &\sim Bin(n_j; q_j) \\ \theta_j &= \text{logit}(q_j) \\ \theta_j \mid \theta &\sim N(\theta, \sigma^2) \\ \theta &\sim N(\theta_0, \sigma_0^2) \\ \sigma^2 &\sim Inv-Gamma(\alpha_0, \beta_0) \end{aligned}$$

The prior construction assumes the logit of response probabilities across cohorts are exchangeable. This shrinks the response probabilities to a common value, θ_0, thus borrowing across cohorts. However, it leads to **the inflate the type I error rate**.

CBHM To solve the problem of inflated type I error, Chu and Yuan[3] proposed a calibrated BHM in which the shrinkage parameter σ^2 is defined as a function of the measure of homogeneity among cohorts. Here, an empirical Bayesian approach is used based on data.

MUCE

Consider a multiple cohort expansion trial that evaluates the efficacy of a new drug of I doses in J indications. Let q_{ij}, n_{ij} and y_{ij} denote the response rate, the number of patients and responds at dose i in indication j, respectively. The objective of the trial to test whether the targeted drug is effective in each dose-indication cohort,

$$H_{0,ij} : q_{ij} \leq p_{0j} \quad \text{versus} \quad H_{1,ij} : q_{ij} > q_{0j}$$

where q_{0j} is the reference rate under which the drug is deemed futile.

We assume y_{ij} follows a hierarchical model,

$$\begin{aligned} y_{ij} \mid n_{ij}, q_{ij} &\sim Bin(n_{ij}, q_{ij}) \\ \theta_{ij} &= \text{logit}(p_{ij}) \\ \theta_{ij} &\sim f_1(\theta_{ij})I\{\lambda_{ij} = 1\} + f_2(\theta_{ij})I\{\lambda_{ij} = 2\}, \\ \lambda_{ij} &= \begin{cases} 1, & Z_{ij} \leq 0 \\ 2, & Z_{ij} \geq 0 \end{cases} \qquad (1) \\ Z_{ij} &\sim N(\xi_i + \eta_j, 1); \;\; \xi_i \sim N(\xi_0, \sigma_\xi^2); \;\; \eta_j \sim N(\eta_0, \sigma_\eta^2) \qquad (2) \end{aligned}$$

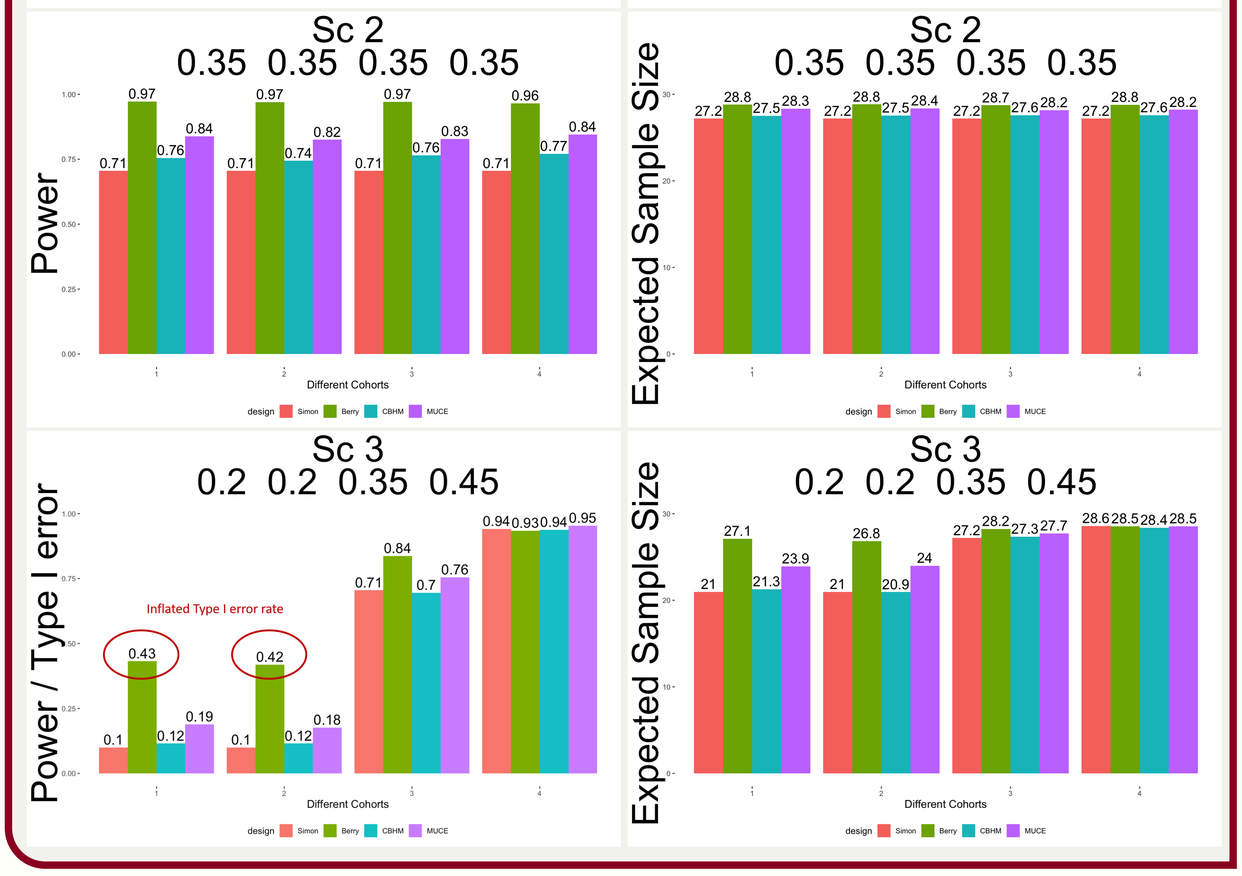
Here, f_1 and f_2 are half-Cauchy truncated at $(-\infty, \text{logit}^{-1}(p_{0j})]$ and $(\text{logit}^{-1}(p_{0j}), \infty)$. In contrast to BHM and CBHM, we **make inference based on λ_{ij} in (1) instead of θ_{ij}**, which helps **handle multiplicity control automatically**. And by **introducing ξ_i and η_j in (2)**, we could **borrow information across doses and indications, even with different degree of borrowing**.

Simulation Studies

Simulation Settings We simulated trials with 1 dose and 4 indications, the reference and target ORRs for each cohort are 0.2 and 0.35, respectively. With the Type I error rate controlled at 0.1 for each cohort, for Simon's 2-stage design, we have $r_1 = 2$, $n_1 = 13$, $r = 8$, $n = 29$ to achieve a 70% power. For Berry's BHM, CBHM and MUCE, the maximum sample size $n = 29$, and the interim analysis for futility are conduct when 10 and 20 subjects have been evaluated.

In BHM and CHBM, the interim futility stopping and the decision of declaring efficacy at the end of the trial are both based on posterior distribution of p_j. In MUCE, they are based on posterior of λ_{ij}.

Simulation Results For Scenarios 1-3



Conclusion

1) MUCE is a smart Bayesian approach with higher power than the traditional independent Simon's 2-stage design.

2) Compared to Berry's method, MUCE decreases the false positive rate when both promising and unpromising cohort arms exist in the study. And compared to CBHM, MUCE has higher power in selecting the promising arms.

3) MUCE can save the sample size if aiming for the same power as the Simon's 2-stage design
   – Under Scenario 2, MUCE has ≥ 0.8 power for each arm with 29 patients per arms, compared to a Simon's 2-stage design with 46 patients per arm that generates 0.8 power.

4) MUCE is capable of dealing with flexibility in borrowing from multiple doses and multiple indications with different reference rates of each cohort.

Reference

[1] FDA draft guidance (2018). Expansion Cohorts: Use in First-In-Human Clinical Trials to Expedite Development of Oncology Drugs and Biologics Guidance for Industry.

[2] Berry, S. M., Broglio, K. R., Groshen, S., & Berry, D. A. (2013). Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. Clinical Trials, 10(5), 720-734.

[3] Chu, Y., & Yuan, Y. (2018). A Bayesian basket trial design using a calibrated Bayesian hierarchical model. Clinical Trials, 15(2), 149-158.