A BAYESIAN DESIGN FOR PLATFORM TRIALS WITH TEMPORAL CHANGES

Chen Zhang

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Biostatistics and Health Data Science
Indiana University

May 2025

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Master's Thesis Committee

Yong Zang, Ph.D., Chair

William F. Fadel, Ph.D.

Pengyue Zhang, Ph.D.

Chen Zhang

A BAYESIAN DESIGN FOR PLATFORM TRIALS WITH TEMPORAL CHANGES

The platform trial, which aims to find the best treatment for a disease by sequentially investigating multiple treatments in a single trial, has become increasingly popular in recent decades. An inherent problem for a platform trial is how to borrow information from the non-current controls to improve the efficiency of the statistical inference. The practical solution of directly combining all the control patients does not work due to the population heterogeneity between the concurrent and non-current controls. The temporal changes are the significant resources for that heterogeneity, which will affect patients' responses over time. In this paper, we develop a Bayesian design to evaluate treatment effects of platform trials accounting for temporal changes. We treat each cohort of patients as a matching set and develop a conditional likelihood method to eliminate the impact of temporal changes. The performance of the proposed method is evaluated through simulation studies.

KEY WORDS: Bayesian modeling; Conditional likelihood; Platform trial; Temporal changes

Yong Zang, Ph.D., Chair

William F. Fadel, Ph.D.

Pengyue Zhang, Ph.D.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Platform trials have drawn tremendous attention due to their great flexibility and efficiency (Mullard, 2014; Berry et al., 2015; Yuan et al., 2016; Ventz et al., 2018; Meyer et al., 2020). It is a dynamic process where experimental agents sequentially enter the trial for evaluation. During the trial, in light of interim data, agents in the trial may be stopped (or graduated) due to futility (or superiority), and new agents may be added to the trial for evaluation. Berry et al. (2015) described platform design as a strategy for evaluating multiple treatments. Yuan et al. (2016) proposed a Bayesian platform design to evaluate multiple single and drug-combination therapies, as well as to adaptively drop and graduate existing treatments and add new treatments during the trial. Saville and Berry (2016) presented a simulation study to explore the efficiencies of platform trial designs relative to a traditional two-arm strategy. Hobbs et al. (2018) used Bayesian predictive probability to evaluate a controlled multi-arm platform trial using multiple comparisons. Kaizer et al. (2018) discussed group sequential methods for platform trials. Tang et al. (2018) proposed a Bayesian platform design to efficiently screen drug combinations using a Bayesian hierarchical model with adaptive shrinkage. Li et al. (2018) proposed a platform design to prioritize the treatment and dose for early phase trials. Some examples of platform trials include the BATTLE trial in advanced non-small-cell lung cancer (Kim et al., 2011; Liu and Lee, 2015), I-SPY2 trial for advanced breast cancer (Barker et al., 2009), NCI-MATCH trials (Mullard, 2015; Barrulet and Matulonis, 2018), Lung-AMP trial (Steuer et al., 2015) and BRAF V600 trial (Kopetz et al., 2020), among others. Platform trials are also used to develop treatments for COVID-19 (Mullard, 2020).

One unique feature that renders platform trials high efficiency is their ability to use a common control to evaluate a pool of investigation agents. The common control serves as the backbone of platform trials. During the trial, in-trial agents may exit, and new agents may enter the trial, whereas the common control arm typically stays throughout the trial. As a result, for a given experimental agent, the control arm data can be divided into concurrent control and non-concurrent control. The perpetuity of the control arm removes the requirement of a separate control for each agent, thereby improving trial efficiency, but as a tradeoff, it makes the comparison of experimental agents with the control more vulnerable to temporal changes (i.e., time trend). Temporal changes may be caused by

various factors, including but are not limited to the changes in eligible criteria, study sites, patient population, or standard of care (Friedman et al., 2010). In the presence of temporal changes, non-concurrent control is systematically different from the concurrent control, and thus the standard approach of comparing the treatment arm with the control arm in the platform trial results in biased estimate of the treatment effect and inflated type I errors.

In this article, we show that in the presence of temporal changes, even using only concurrent control, the standard methods for estimating and testing of the treatment effect of the experimental agent versus the control can be biased. We propose a conditional likelihood approach to address the issue. The proposed method is robust without making any parametric assumption on the temporal changes, and we show that the resulting estimate has causal interpretation under standard causal inference conditions.

Conditional Likelihood Method

Suppose a total of K experimental arms have sequentially entered a platform trial to be evaluated against a common control. Let $k = 0, 1, \cdots K$ index the treatment arms, with $k = 0$ denoting the control, and denote the initiation of the trial as the time 0. The control arm starts at time 0 and stays "active" until the completion of the platform trial (denoted as time U). We use "active" to denote that an arm is ongoing and eligible to receive patients, and use "effective" to denote that an arm is efficacious for treating patients. The K experimental arms enter the trial at different time points, and by the interim time τ, some of them may have been completed and exited the trial (e.g., stop early due to futility or superiority, or reach the pre-specified maximum sample N_k). As a result, the set of active experimental arms may vary from one time point to another, depending on whether some experimental arms are completed or new experimental arms are added. Throughout the trial, patients are randomized among the control and active treatment arms. The primary objective is to evaluate the efficacy of each experimental arm with respect to the control. We focus on a binary efficacy endpoint (e.g., objective response). Extension to a normal endpoint (e.g., pharmacodynamics or biomarker endpoint) is straightforward. In what follows, we first consider the comparison of the experimental arm k to its concurrent control only and then extend the method to incorporate the non-concurrent control.

Suppose during the time period of (τ_{0k}, τ_{1k}), n_0 and n_k patients are (concurrently) randomized into the control and the kth experimental arm. Let $t_{ik} \in (0, U)$ be the randomization time in the whole platform trial for the ith patient among the $n_0 + n_k$ subjects, and $Z_{ik} = 0$ or k be the corresponding treatment option for $i = 1, \cdots, n_0 + n_k$. We investigate the binary outcome in this paper and define $Y_{ik} = 0,1$ as the response outcome. We use logistic regression to fit the data and get

$$\text{logit}\,(Y_{ik} = 1) = \alpha(t_{ik}) + \beta_k \,\text{I}\,(Z_{ik} = k), \qquad (1)$$

where $\alpha(t_{ik})$ is the time-varying intercept due to temporal changes and β_k is the treatment effect.

Although only the concurrently randomized controls are used in model (1), it can be shown that the randomization cannot eliminate the temporal effects. To see this, when the conventional logistic regression is used under the assumption that the temporal effect is controlled due to randomization, the maximum likelihood estimate $\widetilde{\beta_k}$ based on model (1) can be expressed as

$$\widetilde{\beta_k} = \text{Log}\left(\frac{\left\{\sum_{i=1}^{n_0+n_k} I(Z_i=k)I(Y_{ik}=1)\right\}\left\{\sum_{i=1}^{n_0+n_k} I(Z_i=0)I(Y_{ik}=0)\right\}}{\left\{\sum_{i=1}^{n_0+n_k} I(Z_i=k)I(Y_{ik}=0)\right\}\left\{\sum_{i=1}^{n_0+n_k} I(Z_i=0)I(Y_{ik}=1)\right\}}\right)$$

Then, let $\varphi_{ik} = \Pr(Z_{ik} = k)$, given $\alpha(t_{ik}) \neq \alpha(t_{i'k})$ for any $t_i \neq t_{i'}$, we get

$$\widetilde{\beta_k} \to \text{Log}\left(\frac{\left\{\sum_{i=1}^{n_0+n_k} \phi_{ik}\frac{e^{\alpha(t_{ik})+\beta_k}}{1+e^{\alpha(t_{ik})+\beta_k}}\right\}\left\{\sum_{i=1}^{n_0+n_k}(1-\phi_{ik})\frac{1}{1+e^{\alpha(t_{ik})}}\right\}}{\left\{\sum_{i=1}^{n_0+n_k} \phi_{ik}\frac{1}{1+e^{\alpha(t_{ik})+\beta_k}}\right\}\left\{\sum_{i=1}^{n_0+n_k}(1-\phi_{ik})\frac{e^{\alpha(t_{ik})}}{1+e^{\alpha(t_{ik})}}\right\}}\right) \neq \beta_k$$

Therefore, the conventional logistic regression does not work, and additional effort is needed to eliminate the impact of $\alpha(t_{ik})$.

We propose to control the temporal changes using the conditional likelihood method. In particular, the block randomization is generally used in RCTs to randomize a cohort of patients during a short period of time. Therefore, it is reasonable to speculate that patients in the same cohort share approximately the same temporal effect. Hence, if we apply the conditional likelihood method to each cohort of patients, we should be able to control the temporal effects by eliminating them from the likelihood function. In particular, we define L_k as the total cohorts of patients who have been sequentially randomized between treatment k or 0. We also define a cohort size m_l for the lth cohort of patients with the corresponding randomization times $t_{\sum_{j=0}^{l-1} m_j+1k}, \cdots, t_{\sum_{j=0}^{l} m_j k}$ given $m_0 = 0$ for $l = 1, \cdots, L_k$. Similar to the piecewise exponential models, the randomization times within the same block of patients should be close to each other and share about the same temporal effect. Therefore, we have the following approximation:

$$\alpha\left(t_{\sum_{j=0}^{l-1} m_j+1k}\right) \approx \cdots \approx \alpha\left(t_{\sum_{j=0}^{l} m_j k}\right) \equiv \alpha\left(t_{[l]k}\right)$$

Then, the complete likelihood function for treatment k can be written as

$$L_{1k} = \prod_{l=1}^{L_k} \prod_{i=1}^{m_{lk}} \frac{\left(e^{\alpha(t_{[l]k})+\beta_k \mathrm{I}(Z_{ik}=k)}\right)^{\mathrm{I}(Y_{ik}=1)}}{1+e^{\alpha(t_{[l]k})+\beta_k \mathrm{I}(Z_{ik}=k)}} \tag{2}$$

The complete likelihood function (2) involves nuisance parameter $\alpha(t_{[l]k})$ for $l = 1, \cdots, L_k$, which are difficult to estimate when the number of cohort L_k is large. We propose to use the conditional likelihood method to eliminate the nuisance parameters. Let $n_{l0} = \sum_{i=1}^{m_l} \mathrm{I}(Z_{ik} = 0)$ and $n_{lk} = \sum_{i=1}^{m_l} \mathrm{I}(Z_{ik} = k)$ be the number of patients randomized to treatment 0 or k in the lth cohort. Let $r_{l0} = \sum_{i=1}^{m_l} I(Y_{ik} = 1)I(Z_{ik} = 0)$ and $r_{lk} = \sum_{i=1}^{m_l} \mathrm{I}(Y_{ik} = 1)\mathrm{I}(Z_{ik} = k)$ be the number of patients responding favorable to treatment 0 or k in the lth cohort and define $r_l = r_{l0} + r_{lk}$. Let X_{l0} and X_{lk} be the random variables for the number of responses in treatment 0 or k and define $X_l = X_{l0} + X_{lk}$. Then, after fixing $X_l = r_l$, the conditional likelihood function can be written as:

$$\begin{aligned} \boldsymbol{L_{2k}} &= \prod_{l=1}^{L} \Pr(X_{l0} = r_{l0}, X_{lk} = r_{lk} | X_l = r_l) \\ &= \prod_{l=1}^{L} \frac{C_{n_{l0}}^{r_{l0}} \frac{e^{\alpha(t_{[l]k}) r_{l0}}}{\left(1+e^{\alpha(t_{[l]k})}\right)^{n_{l0}}} C_{n_{lk}}^{r_{lk}} \frac{\left(e^{\alpha(t_{[l]k})+\beta_k}\right)^{r_{lk}}}{\left(1+e^{\alpha(t_{[l]k})+\beta_k}\right)^{n_{lk}}}}{\sum_{(r_{l0}^*, r_{lk}^*) \in W_{r_l}} C_{n_{l0}}^{r_{l0}^*} \frac{e^{\alpha(t_{[l]k}) r_{l0}^*}}{\left(1+e^{\alpha(t_{[l]k})}\right)^{n_{l0}}} C_{n_{lk}}^{r_{lk}^*} \frac{\left(e^{\alpha(t_{[l]k})+\beta_k}\right)^{r_{lk}^*}}{\left(1+e^{\alpha(t_{[l]k})+\beta_k}\right)^{n_{lk}}}} \\ &= \prod_{l=1}^{L} \frac{C_{n_{l0}}^{r_{l0}} C_{n_{lk}}^{r_{lk}} e^{\beta_k r_{lk}}}{\sum_{(r_{l0}^*, r_{lk}^*) \in W_{r_l}} C_{n_{l0}}^{r_{l0}^*} C_{n_{lk}}^{r_{lk}^*} e^{\beta_k r_{lk}^*}}, \end{aligned} \tag{3}$$

where W_{rl} contains all the possible combinations of (r_{l0}^*, r_{lk}^*) satisfying $r_{l0}^* + r_{lk}^* = r_l$. Noticing that function (3) only involves the treatment effect β_k. Hence, by using this conditional likelihood function, we should be able to appropriately adjust the temporal effects $\alpha(t_i)$. Because this method uses only the concurrent controls, we refer to it as the concurrent-only method.

Simulation Studies

We conducted comprehensive simulation studies to investigate the performance of the proposed concurrent-only method. We compared the proposed method with the pool method which completely ignores the temporal changes and pools the concurrent and non-current arms together.

We simulated the platform trial illustrated in Figure (1). We assume that the temporal effect changes only when the treatment arms change and let $p_0 = (p_{01}, p_{02}, p_{03})$ be the response rates for the control arm in different sub-trials. We used equal randomization to randomize each cohort size of patients to different treatment arms and let the cohort size equals the number of treatment arms at the randomization time. We defined n_1, and n_2 and n_3 be the number of cohorts randomized at the three sub-trials and specified $n_1 = 50$, $n_2 = 25$ and $n_3 = 25$ through our simulation studies. We also fixed the odds ratio for the treatment effect of T_1 as $\beta_1 = 0.5$ and focused on the evaluation of β_2, the odds ratio for the treatment effect of T_2. We assigned β_2 a non-informative Beta prior Beta (0.5,0.5) and reported the operating characteristics of different methods in evaluating β_2, including bias, 95% coverage, mean-square error (M.S.E.), and type I error or power. As a Bayesian method, we defined that the study success will be claimed if Pr ($\beta_2 > 0$|data) > 0.95. All the results were calculated based on 10,000 replicates.

Table 1 shows the simulation results. Scenarios 1 and 2 are dedicated to the null hypothesis with $\beta_2 = 0$. There is no temporal effect in scenario 1 with $p_{01} = p_{02} = p_{03} = 0.2$. All the methods under comparison yield good performances under scenario 1. Every method reports a type I error around 0.05, a 95% coverage around 0.95, and an acceptable bias. In terms of the mean-square error comparison, the pool method utilizes all the information. It yields the smallest value of 0.224. The concurrent-only method is the conservative method and therefore yields the large value of 0.374. In scenario 2, the control arm's response rate increases from 0.4 in the first sub-trial to 0.6 in the second sub-trial and then keeps consistent in the third sub-trial, so the temporal changes exist. The Pool method fails to control the bias and type I error in the presence of temporal changes. The Pool method reports a huge bias of 0.424 with an unacceptable type I error above 0.3. The concurrent-only method and BRM can well control the bias and type I error. For 95%

coverage, the concurrent-only method yields a value around 0.95. The pool method performs worse with a substantially lower value of 0.777.

Scenarios 3 to 10 are dedicated to the alternative hypothesis with different temporal change patterns and β_2 fixed at 1. In scenario 3, the control arm's response rate monotonically decreases from 0.4 in the first sub-trial to 0.3 and 0.1 in the second and third sub-trials. The concurrent-only method yields the best performances among all the categories under comparison. The pool method failed to adjust the temporal changes and dramatically underestimate the treatment effect (bias -0.621 and -0.227). The results from scenarios 4 to 9 are similar to those in scenario 3. There are no temporal changes in the last scenario. The pool method is served as the golden standard in this scenario and performs slightly better than the concurrent-only method. In general, based on the simulation results in Table 1, the concurrent-only method is undoubtedly a better method than the pool method accounting for the temporal changes.

Conclusion

In this paper, we propose a novel Bayesian method to evaluate the treatment effects of the platform trial in the presence of temporal changes. We treat each cohort of patients under randomization as a matching set and construct a conditional likelihood function to estimate the treatment effects, which treats the temporal effects as nuisance parameters and eliminates them from the estimation process. Simulation studies show that the proposed method yields desirable operating characteristics and outperforms the existing method.

References

Barker AD, Sigman CC, Kelloff GJ, et al. I-spy 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical Pharmacology & Therapeutics 2009;86:97-100.*

Barroilhet L, Matulonis U. The NCI-MATCH trial and precision medicine in gynecologic cancers. *Gynecologic Oncology* 2018; 148: 585-590.

Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. *Journal of the American Medical Association* 2015; 313: 1619-1620.

Friedman LM, Furberg C, DeMets DL, Reboussin DM, Granger CB, et al. *Fundamentals of clinical trials*, volume 4. 2010. Springer.

Hobbs BP, Chen N, Lee JJ. Controlled multi-arm platform design using predictive probability. *Statistical Method in Medical Research* 2018; 27; 65-78.

Kaizer AM, Hobbs BP, Koopmeiners JS. *A Multi-source adaptive platform design for testing sequential combinatorial therapeutic strategies*. Biometrics 2018; 74: 1082-1094.

Kim ES, Herbst RS, Wistuba II, et al. The BATTLE trial: personalizing therapy for lung cancer. *Cancer Discovery* 2011; 1: 44-53.

Kopetz S, Grothey A, Cutsem EV, et al. Encorafenib plus cetuximab with or without binimetinib for BRAF V600E-mutant metastatic colorectal cancer: Quality-of-life results from a randomized, three-arm, phase III study versus the choice of either irinotecan or FOLFIRI plus cetuximab (BEACON CRC). *Journal of Clinical Oncology* 2020; 38: suppl.4039.

Li Y, Wang M, Cheung YK. Treatment and dose prioritization in early phase platform trials of targeted cancer therapies. *Journal of the Royal Statistical Society*, Series C 2018; 68: 475-491.

Liu S, Lee JJ. An overview of the design and conduct of the BATTLE trials. *Chinese Clinical Oncology* 2015; 4: 3.

Meyer EL, Mesenbrink P, Dunger-Baldauf C, Fulle H, Glimm E, Li Y, Posch M, Konig F. The evolution of master protocol clinical trial designs: a systematic literature review *Clinical Therapeutics* 2020; 42:1330-1360.

Mullard A. Multicompany trials adapt todisciplines beyond cancer. *Nature Medicine* 2014; 20:3.

Mullard A. NCI-MATCH trial pushes cancer umbrella trial paradigm. *Nature Reviews Drug Discovery* 2015; 14: 513-515.

Mullard A. COVID-19 platform trial delivers. Nature Reviews Drug Discovery 2020; 19: 501.

Saville BR, Berry SM. Efficiencies of platform clinical trials: a vision of the future. *Clinical Trials* 2016; 13: 358-366.

Steuer CE, Papadimitrakopoulou V, Herbst RS, et al. Innovative Clinical Trials: The LUNGMAP Study. *Clinical Pharmacology & Therapeutics* 2015; 97: 488-491.

Tang R, Shen J, Yuan Y. ComPAS: A Bayesian drug combination platform trial design with adaptive shrinkage. *Statistics in Medicine* 2018; 38; 1120-1134.

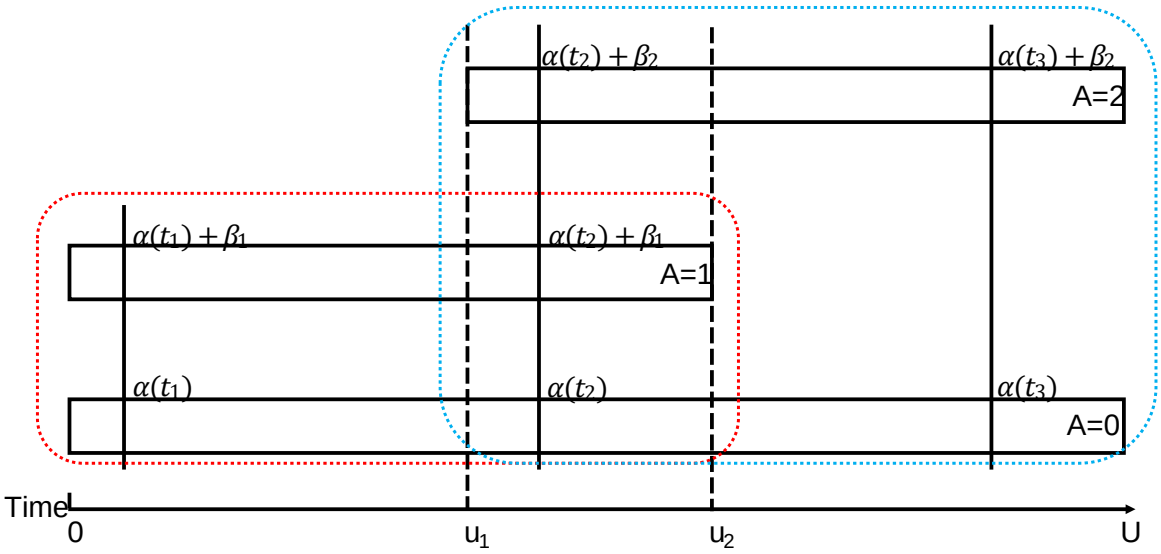
Ventz S, Cellamare M, Parmigiani G, Trippa L. Adding experimental arms to platform clinical trials: randomization procedures and interim analyses. *Biostatistics* 2018; 19: 199-215.

Yuan Y, Guo B, Munsell M, Lu K, Jazaeri A. MIDAS: A Practical Bayesian Design for Platform Trials with Molecularly Targeted Agents. *Statistics in Medicine* 2016; 35: 3892-

Table 1: Operating characteristics of the Concurrent-only (Con.) and Pool methods in evaluating β_2. $n_1 = 50$, $n_2 = 25$, $n_3 = 25$ and $\beta_1 = 0.5$. Equal randomization is used for each block of patients with block size equaling the number of treatment arms.

Method	Bias	95% Coverage	M.S.E	Type I error/ Power
Scenario 1. $\beta_2 = 0$*;* $p_0 = (0.2, 0.2, 0.2)$				
Con.	0.004	0.945	0.374	0.052
Pool	-0.034	0.949	0.224	0.050
Scenario 2. $\beta_2 = 0$*;* $p_0 = (0.4, 0.6, 0.6)$				
Con.	0.000	0.940	0.209	0.054
Pool	0.424	0.777	0.309	0.317
Scenario 3. $\beta_2 = 1$*;* $p_0 = (0.4, 0.3, 0.1)$				
Con.	0.118	0.943	0.615	0.700
Pool	-0.621	0.603	0.519	0.264
Scenario 4. $\beta_2 = 1$*;* $p_0 = (0.4, 0.1, 0.3)$				
Con.	0.111	0.942	0.539	0.695
Pool	-0.621	0.607	0.517	0.263
Scenario 5. $\beta_2 = 1$*;* $p_0 = (0.5, 0.6, 0.2)$				
Con.	0.102	0.946	0.447	0.738
Pool	-0.359	0.835	0.242	0.571
Scenario 6. $\beta_2 = 1$*;* $p_0 = (0.4, 0.2, 0.2)$				
Con.	0.098	0.937	0.472	0.736
Pool	-0.540	0.684	0.428	0.346
Scenario 7. $\beta_2 = 1$*;* $p_0 = (0.5, 0.4, 0.4)$				
Con.	0.092	0.934	0.314	0.804
Pool	-0.169	0.917	0.166	0.742
Scenario 8. $\beta_2 = 1$*;* $p_0 = (0.6, 0.4, 0.4)$				
Con.	0.091	0.934	0.309	0.803
Pool	-0.373	0.810	0.278	0.533
Scenario 9. $\beta_2 = 1$*;* $p_0 = (0.6, 0.6, 0.2)$				
Con.	0.099	0.947	0.466	0.728
Pool	-0.568	0.649	0.436	0.316
Scenario 10. $\beta_2 = 1$*;* $p_0 = (0.2, 0.2, 0.2)$				
Con.	0.112	0.945	0.486	0.749
Pool	0.022	0.945	0.159	0.834

Figure 1: An illustrative example of platform trial



Curriculum Vitae

Chen Zhang

Education

Indiana University at Indianapolis *Master Degree of Science - Biostatistics,*
-Department of Biostatistics and Health Data Science
-May 2025

- Course Projects
  - Employ simple and multiple linear methods to find the relationship between birthweight and maternal factors using SAS
  - Employ simple and multiple logistic regression methods to find the relationship between stroke and risk factors using SAS
  - Employ Kaplan-Meier and the Cox proportional hazards model to identify key prognostic factors impacting Breast Cancer survival rates using R
- Research Experience
  - Randomized Prospective Study Comparison of Erector Spinae Plane Block and Intrathecal Opioid for Postoperative Analgesia after Laparoscopic Colorectal Surgery
  - Work with Professor Ziyue Liu from IUI in reporting statistical analysis according to the study protocol and requirements by using SAS
  - Work with Professor Yong Zang from IUI in developing Bayesian model to test treatment effect from platform trials adjusting population drift
- Publications
  - A Bayesian Design for Platform Trials with Temporal Changes (to be published)