

Narrative Summary of YODA Project # 2016-0960 “Bayesian Methods for Comparative Effectiveness Research”

PI: Gary L. Rosner, ScD (Johns Hopkins University)

Objective: We propose to extend a published Bayesian nonparametric model, the hierarchical Dirichlet process mixture model, in a comparative effectiveness research analysis, inferring a treatment’s effectiveness from multiple randomized clinical trials and possibly observational data. We will also evaluate this model with propensity score adjustment.

Results: We were not able to use the YODA data to which we had access to complete our project. We encountered several obstacles along the way. The main obstacle related to our inability to attract qualified personnel to work on the project over an extended period. An additional obstacle was initial difficulty getting our own C programs and R programs to run on the platform. We had to write our own programs to carry out the proposed analyses efficiently, and it took a while to work out the details to allow us to upload and run these programs.

Mr. Shutong Du, a graduate student at the Johns Hopkins Bloomberg School of Public Health, worked on the project for during the 2020-2021 academic year. Mr. Du looked closely at the data for the eight studies in erythropoietin. He determined that we could possibly use six of these studies (4 oncology studies and 2 studies of critically ill patients) with change in hemoglobin level (from baseline) as the primary endpoint. We note that this measure was not the primary endpoint in each of the randomized trials.

Since Mr. Du was not familiar with C programming, we evaluated publicly available R packages to determine if any would allow us to carry out our proposed analyses. He compared eight programs that are included in the R package DPpackage.¹ This package contains around 50 programs for Bayesian nonparametric inference in the R statistical environment. Mr. Du used simulated data to evaluate the programs. He discovered some errors or inability to work with the data types for covariates in the studies we have available through the YODA Project.

Additionally, we discovered that we require more storage than allocated for our project; the YODA remote desktop was limited to 200 MB. We also needed faster computation to complete the project, since these programs use Markov chain Monte Carlo to generate posterior samples. This process is time consuming and requires a lot of storage. Finally, Mr. Du graduated in May 2021 and began doctoral studies at a different university.

The final blow occurred when Dr. Chenguang Wang, my co-investigator on this project, left Johns Hopkins University for a job in industry this past summer. Without any additional help, I am, unfortunately, unable to complete the project.

Conclusions: N/A

1. Jara A, Hanson TE, Quintana FA, Muller P, Rosner GL. DPpackage: Bayesian Semi- and Nonparametric Modeling in R. *Journal of Statistical Software*. 2011;40(5):1-30. PubMed PMID: WOS:000289228600001.

2016-0960 “Bayesian Methods for Comparative Effectiveness Research”

YODA Project Response:

We thank Dr. Rosner and his team for sharing the obstacles they faced in completing their YODA Project data request #2016-0960. We would like to make one point of clarification to their report, which is that the YODA Project was not notified of the issues the team experienced with the virtual environment (inadequate space and computing power). If we had been aware, we could have worked with the team to overcome them in a way that would have made the research possible.