A comparison of Bayesian meta-analysis methods for rare adverse events

Jinyi Zhou[1]*, Gary Rosner[2], Chenguang Wang[2], and Hwanhee Hong[1]

[1] Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina, USA

[2] The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland, USA

* Degree Program: Master of Biostatistics; email: jz309@duke.edu

Duke University School of Medicine

Introduction

- Meta-analysis is a commonly used statistical technique for combining results from multiple clinical trials.
- Meta-analysis with extremely rare events causes an issue of data sparsity, leading to zero-event trials that may cause extremely skewed distributions of event frequencies and insufficient statistical power to estimate the effect heterogeneity across studies.
- Bayesian meta-analysis models have the advantage of handling the sparsity due to their flexibility and the ability of employing a wide range of prior specifications.
- We compare the performance of 8 Bayesian meta-analysis methods and explore different prior distributions.

Bayesian meta-analysis model

- Likelihood

$$y_{ik} \sim Bin(n_{ik}, p_{ik}),$$

  - $i = 1, \ldots, I$: study
  - $k = 1$ for the control group; $k = 2$ for the treated group
  - y_{ik}, n_{ik}, p_{ik}: the number of events, the number of subjects, the probability of having an event in group k on study i

- **Logistic regression** model
  We consider models under two assumptions:
  (1) constant treatment effect **(CTE)**
  (2) heterogeneous treatment effect **(HTE)**

$$\textbf{CTE-Logit}: logit(p_{ik}) = \mu_i + dI(k = 2)$$
$$\textbf{HTE-Logit}: logit(p_{ik}) = \mu_i + \delta_i I(k = 2),$$

  - $\mu_i \sim N(0, 10^2)$: the study-specific baseline effect
  - $d \sim N(0, 10^2)$ or $N(0, 2.82^2)$: LOR between two groups
  - $\delta_i \sim N(d, \tau^2)$: the study-specified LOR
  - 4 prior distributions for τ, between-study heterogeneity: $Uniform(0, 2)$, $HalfCauchy(0, 0.5)$, $Pareto(0.5, 0.006)$, and $HalfNormal(0, 16)$

- **Arm-based** model

$$logit(p_{ik}) = \theta_k + \eta_{ik},$$

  - $\theta_k \sim N(0, 10^2)$: the k^{th} treatment effect (log odds of treatment k)
  - η_{ik}: random effects allowing heterogeneity of the log odds
  - $(\eta_{i1}, \eta_{i2})^T \sim BVN((0, 0)^T, \Sigma)$, $\Sigma^{-1} \sim Wishart(\Omega, 2)$, $\Omega = \begin{bmatrix} 0 & 6 \\ 6 & 0 \end{bmatrix}$

  Note The choice of Wishart distribution is important and should depend on data. We chose the one that provided a reasonable prior distribution and the smallest WAIC among our candidates.

- **Beta-hyperprior** model
  - **CTE-Beta** assumes that p_{ik} is consistent across studies.

$$p_k \sim Beta(1, 1)$$

  - **HTE-Beta** assumes that p_{ik} varies across studies.

$$p_{ik} \sim Beta(U_k V_k, (1 - U_k) V_k)$$

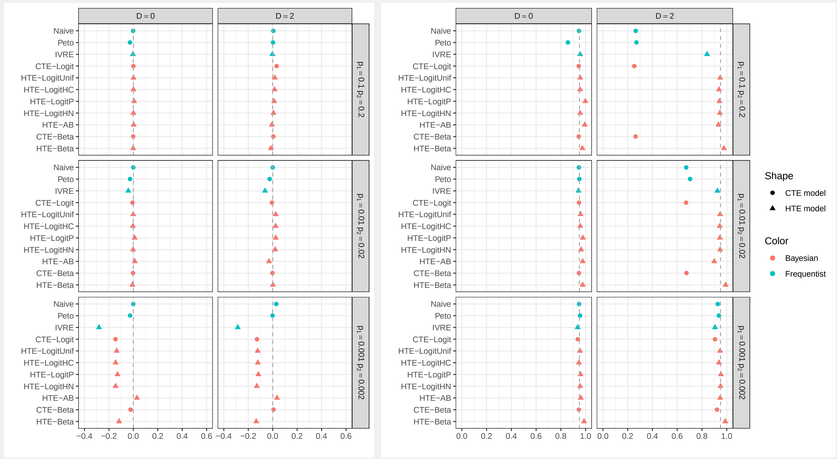
    - $U_k = \frac{a_k}{a_k + b_k} \sim Beta(1, 1)$: the mean of p_{ik}s
    - $V_k = a_k + b_k \sim Gamma^{-1}(1, 0.01)$
    - $\frac{U_k(1-U_k)}{V_k+1}$: study heterogeneity in the probability scale

Simulation study

- **Settings**
  - 1000 simulated meta-analysis data sets
  - Each dataset includes 30 studies
  - $n_{i1} \sim Uniform(50, 1000)$, $n_{i1} = n_{i2}$ (1:1 allocation)
  - Between-study heterogeneity D = 0, 1, 2
  - $p_{ik} \sim Uniform(p_k(1 - 0.5D), p_k(1 + 0.5D))$
  - Three different degrees of event rareness:

(p_1, p_2)	Common	Infrequent	Rare
Null	(0.1, 0.1)	(0.01, 0.01)	(0.001, 0.001)
Alternative	(0.1, 0.2)	(0.01, 0.02)	(0.001, 0.002)

Here, we display the results of the alternative cases when D = 0, 2.



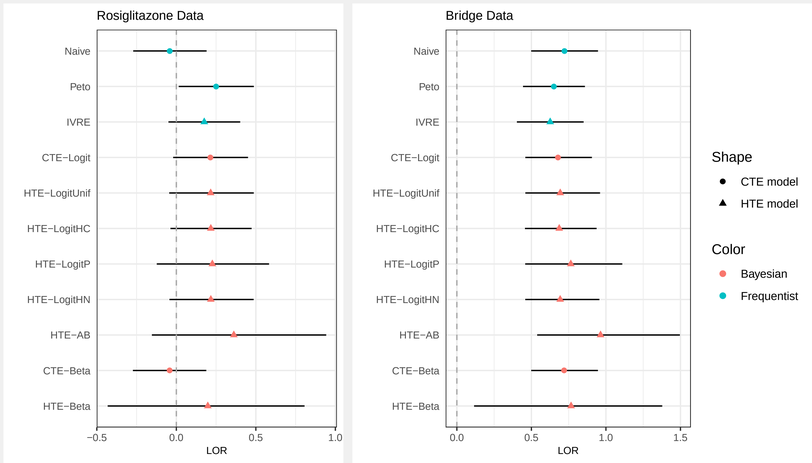
(a) Bias (b) Coverage Probability

- **Results**
  - Bias gets larger as the true risk of event gets smaller. A few methods (Naive, Peto, IVRE, and CTE Bayesian models) provided poor coverage when the outcome is common or infrequent, but widely heterogeneous (D=2).
  - **CTE-Beta** and **HTE-Beta** gave the smallest WAIC when D = 0 and 2, respectively.
  - For the rare outcome case, Peto, HTE-AB, and CTE-Beta models provided unbiased estimates, while Bayesian logistic regression-style models and HTE-Beta provided somewhat biased results.
  - When D = 0, all methods are able to achieve the nominal coverage probability (CP) 0.95, except Peto under the common outcome case.
  - When D = 2, all Bayesian HTE models are able to achieve the nominal CP across all cases. On the contrary, all CTE models (both Bayesian and frequentist) yielded CP lower than 0.95, but these CPs become closer to 0.95 as the true event risk gets smaller.

Real data analysis

- **Rosiglitazone data** study the effect of rosiglitazone on the risk of myocardial infarction (MI).
- **Bridge data** study the effect of the pediatric antidepressant treatment on the risk of suicide attempt.

	Rosiglitazone data	Bridge data
N of Trials	56	27
Pooled risk$_{control}$	0.00848 (136/16022)	0.00756 (112/14811)
Pooled risk$_{treatment}$	0.00815 (159/19509)	0.01544 (256/16578)



- For rosiglitazone data, all methods except Naive and CTE-Beta estimated positive LORs. Only Peto provided a 95% confidence interval excluding 0.
- For Bridge data, all methods provided positive LOR estimates with 95% credible/confident intervals excluding 0.
- For both data, **HTE-AB** gave the smallest WAIC. **HTE-Beta** and **CTE-Logit** gave the second smallest WAICs for Rosiglitazone and Bridge data examples, respectively.

Conclusion

Overall, Bayesian **HTE-AB** with a properly specified Wishart prior and **HTE-Beta** perform well and provide good model fits. Bayesian CTE also performs well with rare events under valid CTE model assumptions. We recommend to fit various Bayesian meta-analysis models and compare the results and model fits.

Future work

To improve the HTE-Beta method, we are implementing a mixture prior using a Dirichlet Process that allows us to employ a weighted Beta prior (between non-informative and informative priors) for U_k.

References

Hong, H., Wang, C., Rosner G. “Meta-Analysis of Rare Adverse Events in Randomized Clinical Trials: Bayesian and Frequentist Methods” 2019 *Clinical Trials*, Forthcoming.