Bayesian Model Averaging-inspired Evidence Synthesis in Federated Healthcare Studies

Shounak Chattopadhyay[1,3], Martijn J. Schuemie[2], Marc A. Suchard[1,3]

[1] Department of Statistics, University of Virginia

[2] Epidemiology, Office of the Chief Medical Officer, Johnson & Johnson, Titusville, NJ

[3] Department of Human Genetics, University of California, Los Angeles

Background

OHDSI network studies involve massive observational healthcare databases in a federated learning setup to protect patient privacy. Existing evidence synthesis approaches across such networks typically assume independence of the unit-level analyses being aggregated, allowing traditional random-effects meta analysis [1] or combining likelihoods [2]. However, this assumption is routinely violated, such as when we aggregate inference across different designs executed on the same underlying data source, or when there is overlap among different data sources. Such dependence leads to inefficient point estimates and uncertainty quantification with existing methodology.

Taking inspiration from Bayesian Model Averaging (BMA), we develop a novel principled probabilistic framework for evidence synthesis that is able to aggregate information across dependent analyses in a federated learning setup. We aggregate local-level power posteriors using weights to yield a global posterior of interest, with the weights carefully chosen according to an entropy-based metric that prioritizes studies with greater information content. Furthermore, if further information regarding the association between each unit-level analysis is available, we provide an improved strategy to choose the weights. We show the superiority of our approach over existing approaches via numerous simulation examples and a real-world application in vaccine safety surveillance.

Methods

We assume a general setup where analyses may be dependent, with different studies carried out across a federated network. Each local study produces a power-posterior of the parameter of interest, with the exponent obtained as to ensure local-level analyses have the same rate of

variance as the full-data analysis (which is unavailable due to federated learning constraints). We aggregate across local analyses by computing study-specific weights that characterize the information gained from updating the prior to the posterior [3]. We found this scheme to work better than naively using marginal likelihoods due to their tendency to downweight studies with greater sample sizes.

When the covariance structure across the local-level maximum likelihood estimators (MLEs) is available or estimable, we extend this framework by introducing optimal weights that minimize the variance of the aggregated posterior mean. This yields a more efficient estimator than the previous choice of weights. In the absence of any such covariance information, the weights revert back to the earlier entropy-based weights.

The running example throughout our methodology is that of a “single target, multiple comparator” setup, where a single target is compared to multiple comparator cohorts in a pairwise manner. For example, to measure the efficacy of a target drug of interest, different local studies might compare the target drug to their placebo/comparator drug of choice. We assume different data generating mechanisms, such as Poisson regression or conditional Poisson regression, allowing derivation of the asymptotic covariance between the MLEs and calculation of optimal weights as described above.

Results

Our simulations demonstrate that the BMA-based approach achieves comparable point estimation accuracy to the gold standard full-data posterior, while providing substantially improved uncertainty quantification compared to other competitors, with more calibrated intervals than the Wasserstein posterior (WASP) [4]. In scenarios with local posteriors reflecting inferences with higher variability, the BMA-based approach appropriately reflects the multimodal global distribution (Figure 1).

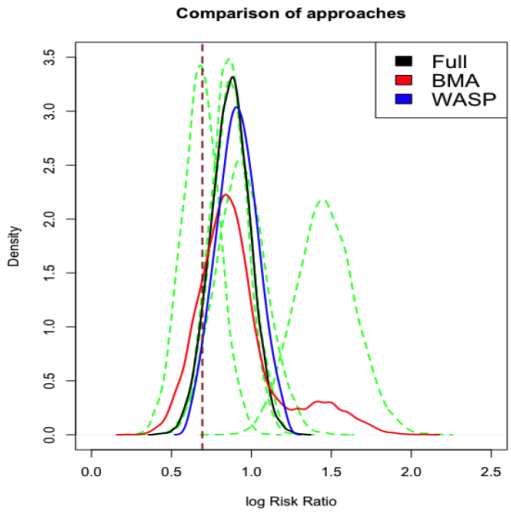

Fig 1: Comparison of full data posterior, BMA-based posterior, and WASP for an example.

As a real-world illustration, we consider the EUMAEUS (Evaluating Use of Methods for Adverse Event Under Surveillance) study [5]. The EUMAEUS study carried out an extensive comparison of different state-of-the-art designs for vaccine safety surveillance leveraging the OHDSI tool-stack on the basis of negative control outcomes. For our aggregation, we considered 15 different variants of the self-controlled case series (SCCS) approach [6], obtained by stratifying analyses by either considering or not considering different factors such as age, sex, and other demographic characteristics. In Figure 2, we compare the WASP with the BMA-based posterior when aggregating risk ratios for “contusion of toenail”, a negative control outcome (with log risk ratio = 0), based on data from the Optum EHR data source. The BMA-based posterior successfully captures the true effect size, while the WASP failed to do so.

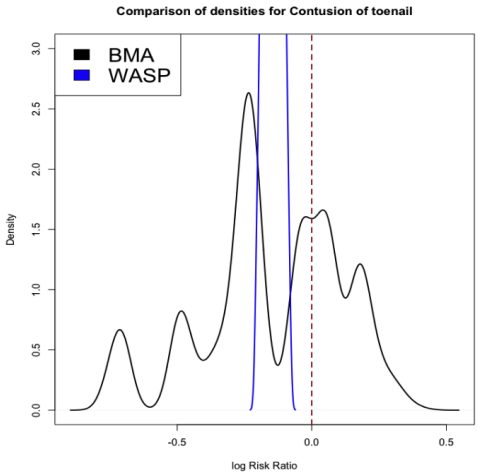

Figure 2: BMA vs WASP posteriors for the real-world null outcome (contusion of toenail) across 15 different SCCS variants.

Conclusion

We propose a generalizable and novel privacy-preserving Bayesian framework for synthesizing evidence across possibly dependent federated studies carried out on massive observational healthcare data. Our approach bridges the gap in meta-analytic methods for dependent unit-level analyses and provides performance improvements over existing approaches that assume independence in settings where such assumptions are violated. The aggregation may be carried out either across dependent data sources or dependent designs and is naturally extended to inferences on higher dimensional parameters.

References

1. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. Bmj. 2011 Feb 10;342.

2. Schuemie MJ et al. Combining Cox regressions across a heterogeneous distributed research network. Stat Methods Med Res. 2022;31(3):438–450.
3. Lindley DV. On a measure of the information provided by an experiment. Ann Math Stat. 1956;27(4):986–1005.
4. Srivastava S, Li C, Dunson DB. Scalable Bayes via barycenter in Wasserstein space. J Mach Learn Res. 2018;19(1):312–346.
5. Schuemie MJ, Arshad F, Pratt N, Nyberg F, Alshammari TM, Hripcsak G, Ryan P, Prieto-Alhambra D, Lai LY, Li X, Fortin S. Vaccine safety surveillance using routinely collected healthcare data—an empirical evaluation of epidemiological designs. Frontiers in Pharmacology. 2022 Jul 6;13:893484.
6. Petersen I, Douglas I, Whitaker H. Self controlled case series methods: an alternative to standard epidemiological study designs. bmj. 2016 Sep 12;354.