Extending Bayesian Inference in Pharmacovigilance Beyond Point Estimates with Massive Parallelization

Trevor R. Shaddox[1] and Marc A. Suchard[1,2,3]

[1]Department of Biomathematics and [2]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles.

[3] Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles.

Introduction

Medical Problem: Adverse drug events (ADEs) are a serious public health risk. Clinical trials lack the sample sizes to detect rare events.

- Vioxx (rofecoxib) has been one of the highest profile cases of a drug with insidious side effects. The cardiac effects of rofecoxib were not appreciated until after the drug had been released to the market.
- This motivates the need for post approval surveillance of drugs, using computational techniques to detect ADEs in repositories of clinical data.

Mathematical Solution: One of the most promising signal detection methods is the Bayesian self-controlled case series (BSCCS).

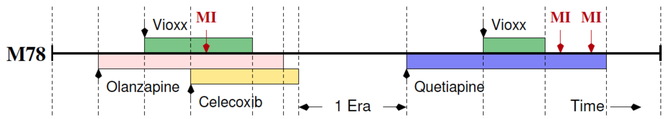

FIGURE 1: BSCCS data model.

- BSCCS assumes adverse events occur as a truncated Poisson processes. $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)'$ are unknown relative risks attributable to each drug.
- Under the model, the number of adverse events in era k of patient i is Poisson distributed. The BSCCS method conditions on the sufficient statistic, the total number of ADEs $n_i = \sum_k y_{ik}$ that a subject experiences over her observation period. The length of exposure is measured in l_{ik}. Taking all subjects as independent yields the log-likelihood as a function of unknown $\boldsymbol{\beta}$.

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{N} \left[\sum_{k=1}^{K_i} \left(y_{ik}\, \mathbf{x}_{ik}'\boldsymbol{\beta} \right) - n_i \log \left(\sum_{k=1}^{K_i} l_{ik}\, e^{\mathbf{x}_{ik}'\boldsymbol{\beta}} \right) \right]$$

- To reduce overfitting, a Normal or Laplace prior probability is placed over each β_j. This reduces the statistical problem to regularized regression, using L_2 and L_1 penalties.
- *Cyclic Coordinate Descent:* Cyclic coordinate descent (CCD) is used to find the maximum *a posteriori* (MAP) estimates $\hat{\boldsymbol{\beta}}_{\boldsymbol{MAP}}$

Beyond Point Estimates:

- Bootstrapping regularized estimators is perilous. (Chatterjee and Lahiri, 2011)
- *Fully Bayesian Implementation:* Extending the existing ideology suggests a fully Bayesian model, with an entire posterior distribution captured.

Goal:

Implement Markov Chain Monte Carlo (MCMC) methods for this problem.

Methods

MCMC:

- The two parameters of interest are the vector of regression coefficients $\boldsymbol{\beta}$ and the hyperprior variance σ of the β_i.
- We address these parameters by constructing the Markov chain with a Metropolis-within-Gibbs sampler.

Compute mode estimates $\hat{\beta}$;
Compute local precision Σ^{-1};
while *generating iterates* **do**
  Select transition kernel;
  if *Metropolis-Hastings* **then**
    Sample $\beta^* \sim \mathcal{N}(\hat{\beta}, \Sigma)$;
    Accept with probability $\alpha = min\left\{\frac{\pi(\beta^*)f(\beta^*)}{\pi(\beta^c)f(\beta^c)}, 1\right\}$;
    if *accept* **then**
      $\beta^c = \beta^*$;
    end
  else
    Perform Gibbs sampling on σ;
    $\sigma \sim \Gamma\left(\alpha + \frac{N}{2}, \kappa + \frac{\sum_i (\beta_i - \mu)^2}{2}\right)$;
    Recompute $\hat{\beta}$;
    Recompute Σ^{-1};
  end
end

Algorithm 1: Outline of Metropolis-within-Gibbs MCMC

Making it work: Video game technology

- Graphics Processor Units (GPUs) offer access to incredible computational performance for parallelizable operations.
- The most expensive task is computing the log-likelihood, a process that is readily amenable to GPU programing.

Results

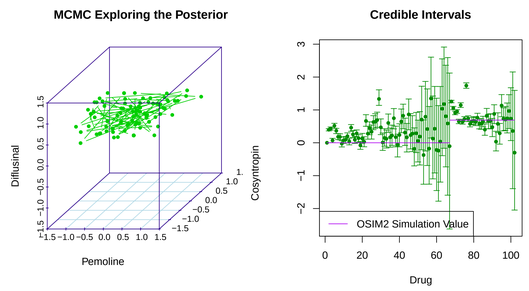

FIGURE 2: Depiction of the MCMC exploring the posterior space and the resulting credible intervals.

OMOP OSIM2 Data: Evaluating the model in the context of known signal

- Coverage matches previous estimates with the OMOP OSIM2 10M MSLR acute myocardial infarction data (condition concept id 500000801).
  - Data: over 350,000 patients exposed to 1256 drugs with over 4,500,000 observations
  - Nominal 95% coverage estimates produce roughly 62% coverage.

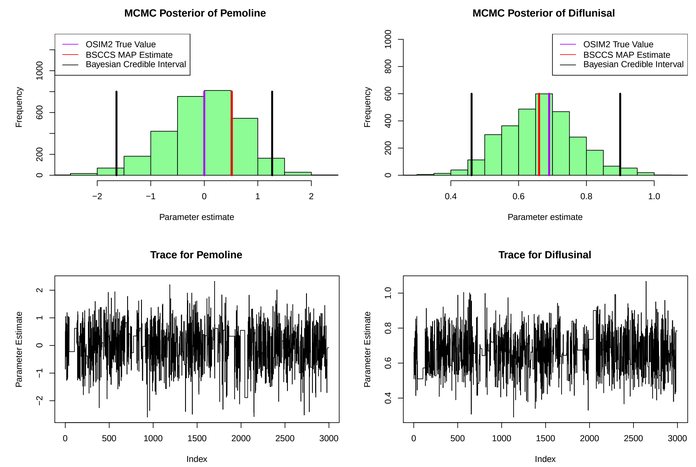

FIGURE 3: Example marginal distributions and traces of the simulations that generated them for two drugs using the OMOP OSIM2 dataset.

Conclusion

Significance:

- Meaningful Bayesian credibile intervals: matching previous estimates of uncertainty around estimated parameter values.
- Reasonable mixing rates: the high dimensionality does not prove an insurmountable obstacle to MCMC implementation in the observational data setting.
- Significant computational speed up: applicability to largest datasets hinges on GPU power.
  - Two-fold speed up at minimum.

Future Work:

- Computational goals:
  - Adaptive kernel for independence sampler.
  - Better exploration of the posterior space.
- Application: compare performance on known signals in real datasets.