Hierarchical expectation propagation for Bayesian aggregation of average data

Andrew Gelman, Columbia University
Sebastian Weber, Novartis

also Bob Carpenter, Daniel Lee, Frédéric Bois, Aki Vehtari, and Amy Racine

23 Sept 2015

Motivation

- Trial design with hierarchical models using patient data and summaries of external data
- More generally: incorporating aggregate data into statistical analysis

Scenario

- Non-inferiority and bio-similarity trials
- Test of a candidate substance against an active control
  - Candidate substance developed in-house: lots of raw data, individual patient level longitudinal data
  - Active control developed externally: we only have data summaries from publications or submission documents
- New trial will be similar to earlier trials of the active control drug, but still to some extent different
- Also the conditions differ in the two experiments

Meta-analysis

- Partial pooling of information from two studies
- Statistical challenges:
  - Only summary data are available on one study
  - Nonlinear model, so no closed-form model for average data
  - Only 2 "groups" so hierarchical modeling is not so easy, requires strong prior information

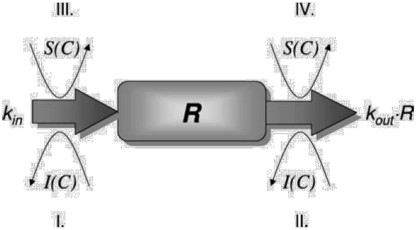
Drug disease modeling of drug responses

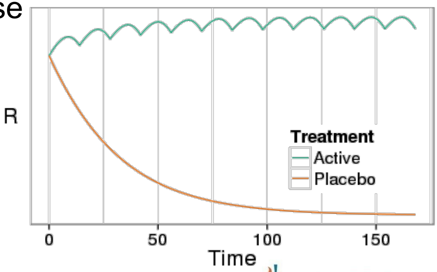
- Key application is clinical trial simulations for study design
- Simulation of drug responses of patients over time
  - New designs can be considered
  - Different endpoints can be explored including time to event
- Hierarchical (population) based models require patient-level data
- Problem: Same disease progression, same patient population and likely similar mechanism of action, but population model describes only in-house drug
- How to learn from published summaries of longitudinal data in the context of nonlinear hierarchical models?

Semi-Mechanistic Turn-Over Models

Linking Pharmacokinetics (PK) with Pharmacodynamics (PD)

- PD response «R» can be safety or efficacy related driven by PK effect on «bio-compartement»
  - Zero order «production» / first order «elimination» of response R
  - 4 variants: zero / first order inhibition / stimulation due to PK
  - Drug response with respect to reference state (placebo)
- Some regimens may lead to PK causing oscillations and hence oscillations in response





Source: Peletie LA et al.; J Pharmacokinet Pharmacodyn. 2005

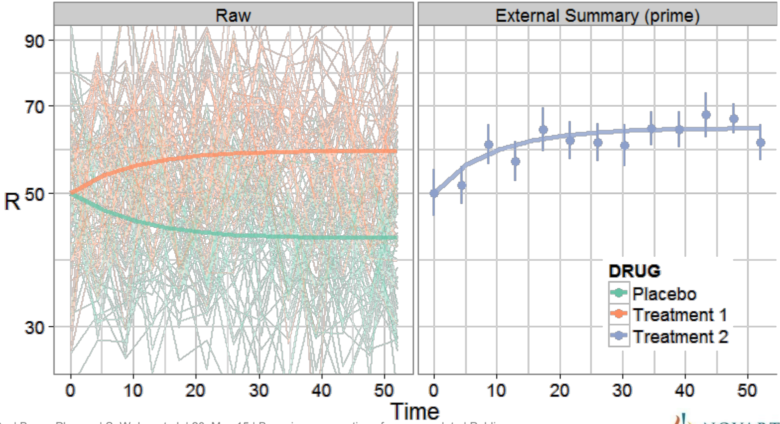
General formulation

- Direct data $y = (y_{jt}; j = 1, \ldots, J; t = 1, \ldots, T)$ with model $p(y_j | \alpha_j, \phi)$
- Hierarchical model $p(\alpha, \phi) = p(\phi) \prod_{j=1}^{J} p(\alpha_j | \phi)$
- Use Stan to draw posterior simulations from $p(\alpha, \phi | y) \propto p(\phi) \prod_{j=1}^{J} p(\alpha_j | \phi) \prod_{j=1}^{J} p(y_j | \alpha_j, \phi)$
- External dataset $y' = (y'_{jt}; j = 1, \ldots, J'; t = 1, \ldots, T')$
- Observe time series of averages: $\bar{y}' = (\bar{y}'_1, \ldots, \bar{y}'_T)$
- External dataset has parameters $\phi' = \phi + \delta$
- Informative prior on δ

Simulated Example Data Set

50 Patients per Treatment Arm Placebo, Treatment 1 & 2

- Solid line is true population mean
- No population differences
- No between-trial variation



NOVARTIS

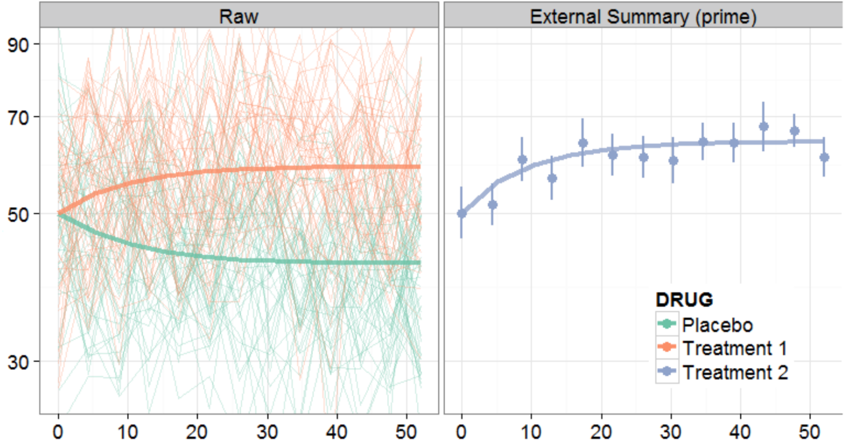
Integrating external summary data

- Simplest case: 2 datasets follow the same model
  - Populations must be comparable
  - Parameters must be identical
  - Natural disease progression must be the same
- Relax these assumptions: $\phi' = \phi + \delta$
- Extreme cases
  - $\delta \equiv 0$: complete pooling
  - $p(\delta) \propto 1$: no pooling

Natural Bayesian approach doesn't work

- Consider external data as latent variables
- y' not observed; all we see is $\bar{y}'$
- Computationally expensive
  - For example, 300 patients and 15 measurements per patient
- Instead, we'll model $\bar{y}'$ directly
- Take advantage of central limit theorem

Multivariate normal approx to $p(\bar{y}'|\phi')$



- Given ϕ' simulate data from 1000 hypothetical patients
- Compute mean M and $T \times T$ covariance matrix S
- Approx $p(\bar{y}'|\phi')$ by $\mathrm{N}(\bar{y}'|M, S/J')$
- J' = number of patients in external data, not the same as the "1000"

Importance sampling algorithm

- Fit model to direct data; get draws from $p(\alpha, \phi | y)$
- For each draw of α, ϕ:
  - Draw δ from prior $p(\delta)$
  - Compute $\phi' = \phi + \delta$
  - Simulate data from 1000 hypothetical patients
  - Approx $p(\bar{y}' | \phi')$ by $\mathrm{N}(\bar{y}' | M, S/J')$
  - Compute importance ratio $\mathrm{N}(\bar{y}' | M, S/J')$

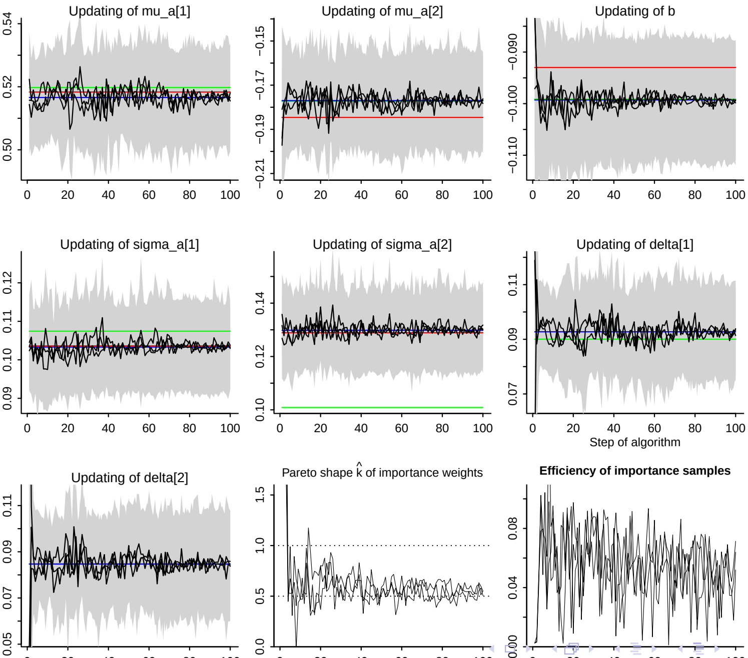
Hierarchical expectation propagation (EP) algorithm

- Simple importance sampling won't work if $p(\delta)$ is broad
- Need iterative algorithm
- At each step, sample from "pseudo-prior" $g(\phi, \delta)$
- Multiply importance ratios by $\frac{p(\phi,\delta)}{g(\phi,\delta)}$
- EP: match moments to get update for g
- Use smoothed importance weights and stable moment matching

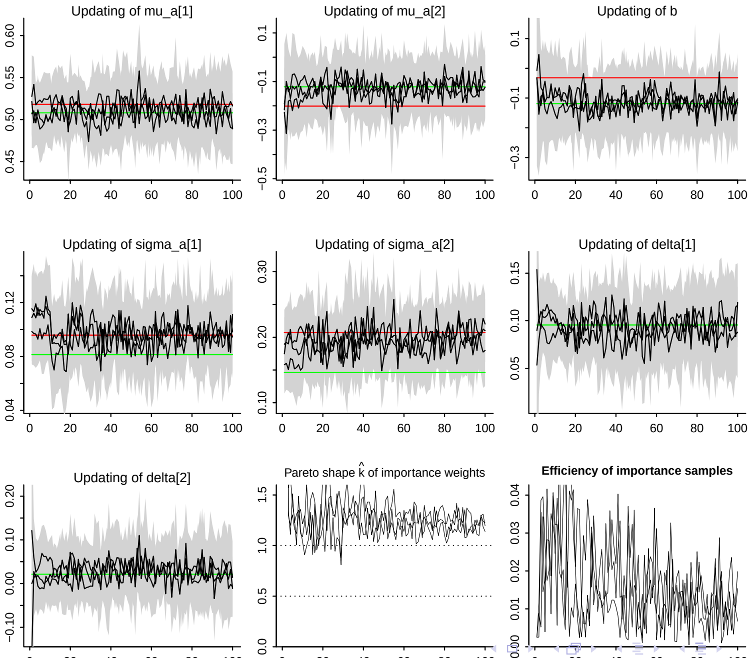
Examples

- Hierarchical linear model:
  - Local data: $y_{jt} \sim \mathrm{N}(\alpha_{j1} + \alpha_{j2}x_t + \beta x_t^2, \sigma_y^2)$
  - $\alpha_j \sim \mathrm{N}(\mu_\alpha, \Sigma_\alpha)$
  - External data: $y_{jt} \sim \mathrm{N}(\alpha'_{j1} + \alpha'_{j2}x_t + \beta x_t^2, \sigma_y^2)$
  - $\alpha'_j \sim \mathrm{N}(\mu'_\alpha, \Sigma_\alpha)$.
  - $\delta = \mu'_\alpha - \mu_\alpha$.
- Hierarchical logistic
- Hierarchical PKPD

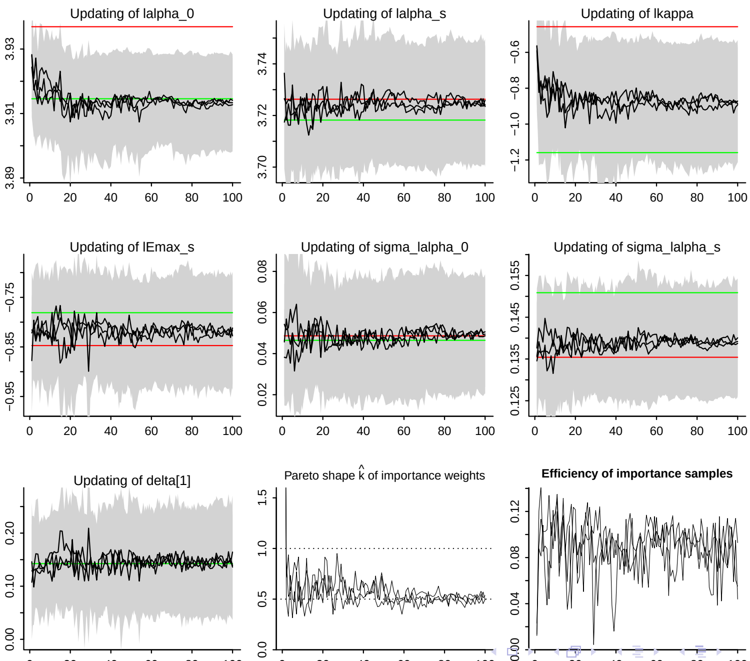
Hierarchical linear model example: Posterior mean +/− sd from EP algorithm from 3 starting points
(Red lines show estimate from local data, blue includes aggregate data, green uses complete data)



Hierarchical logistic example: Posterior mean +/− sd from EP algorithm from 3 starting points
(Red lines show estimate from local data, green uses complete data)



Hierarchical PK/PD example: Posterior mean +/− sd from EP algorithm from 3 starting points
(Red lines show estimate from local data, green uses complete data)



Discussion

- Use fake-data simulation to build trust in results
- Hope to improve efficiency of EP by approximating $p(\phi, \delta)$ (currently approximating $p(\phi)$ and $p(\delta)$)
- Shift parameter δ instead of "cut"
- New way to think about meta-analysis