Bayesian Model Averaging of Longitudinal Dose-Response Models

Richard Payne, PhD

Lilly

Collaborators

- Pallavi Ray
- Mitch Thomann

Dose Response Modeling

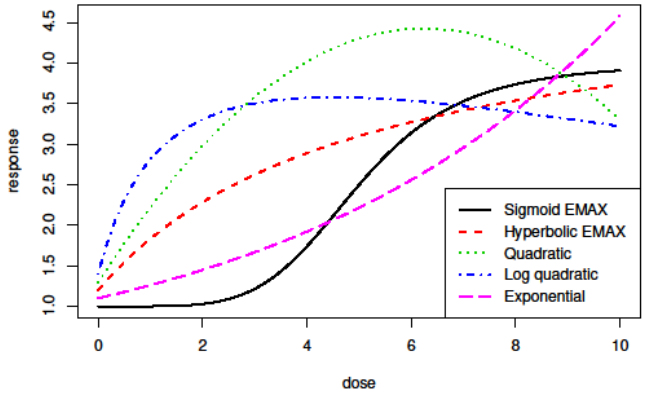
- Why is it important?
  - 15% of failed first-time applications for New Molecular Entities to the FDA were related to uncertainties in dose selection (Sacks 2014).
  - Increasing the number of doses may not improve power, but it provides much greater information on the dose-response curve.
  - Helpful when choosing and/or justifying a dose to regulatory bodies.

Dose Response Models

- What are they?
  - A model which assumes a (non-)parametric form across the dose range, so information across all arms is shared for estimation.
  - Different than pairwise tests which compare doses independently.
  - Examples:

    EMAX (sigmoid and hyperbolic), Quadratic, Exponential, Beta, Linear, Log-linear, Log-quadratic…

With limited data, how does one choose a dose-response model *a priori*?



Choosing a Dose-Response Model

- If there is prior data or information about the dose-response curve (e.g. other molecules in the same class and/or indication) or something is known about the dose-response curve (e.g. monotonic).
  - A suitable dose-response model might be able to be chosen *a priori.*
- How many doses? Two, three, more?
  - More flexible models will require more doses for suitable estimation (e.g. EMAX).
- What if there's not much data? What if there's a plausible scientific hypothesis of non-monotonicity, but it's unclear at which dose this might occur?
  - This is a good candidate for using Bayesian model averaging.

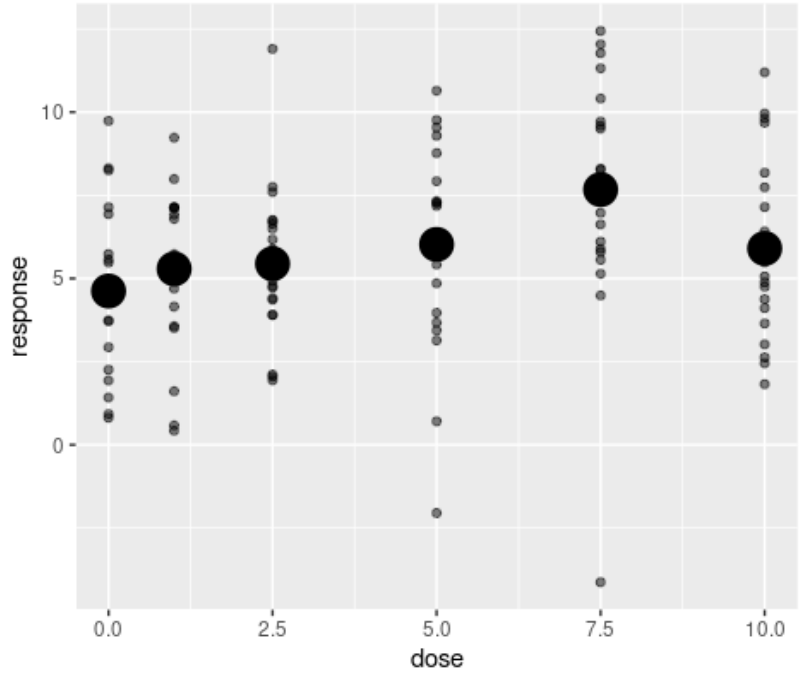
Hypothetical Scenario

- Suppose there's a new molecule with limited phase I data.
- There is no data from other molecules in the same class.
- There is a plausible scientific hypothesis that the dose-response is non-monotonic at high enough doses, but the dose range proposed in the phase 2 dose-finding study is not believed to be in that zone.
- To account for the possibility of non-monotonicity, Bayesian model averaging is used with prior weight of .75 on an EMAX model, and .25 on a quadratic model.

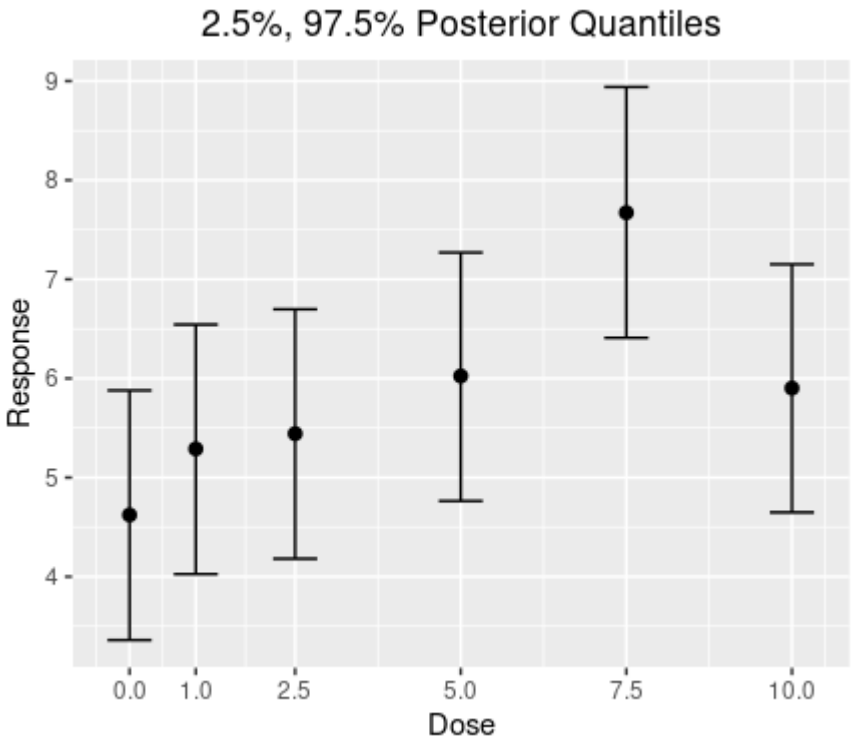
Example

Pairwise p-value (vs placebo) is significant for dose 7.5 (0.005) but not dose 10 (0.14).

Simulated Data

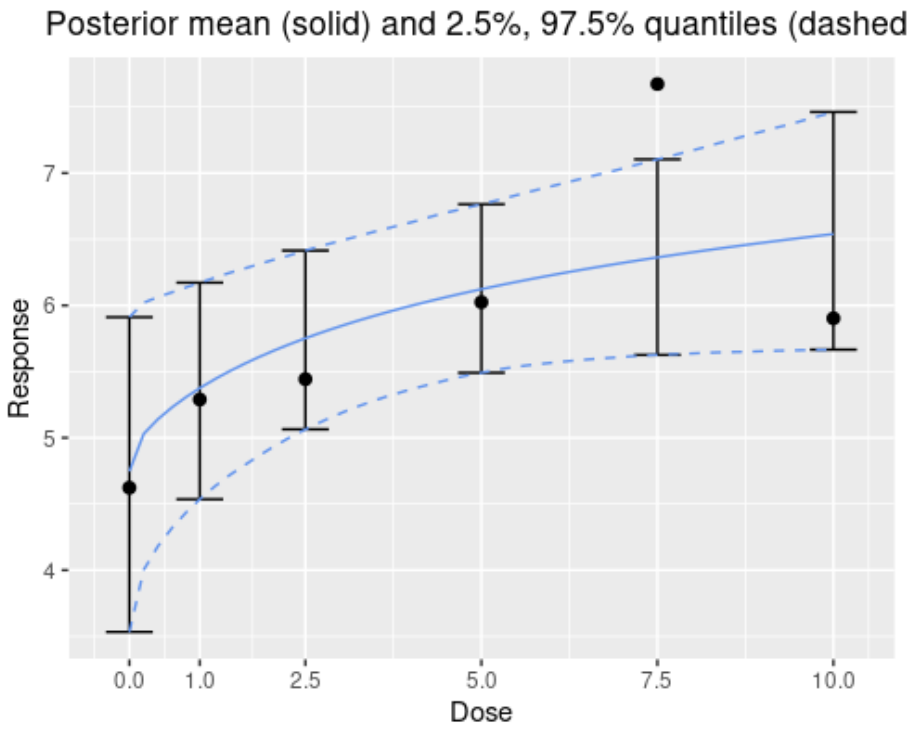


Independent Bayesian Credible intervals

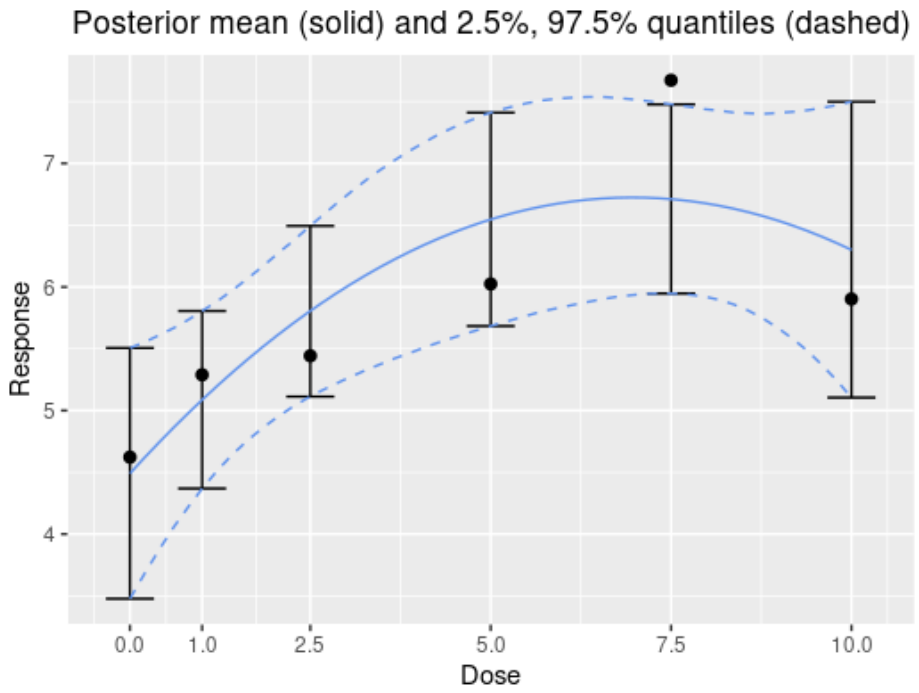


Example: Dose Response Models

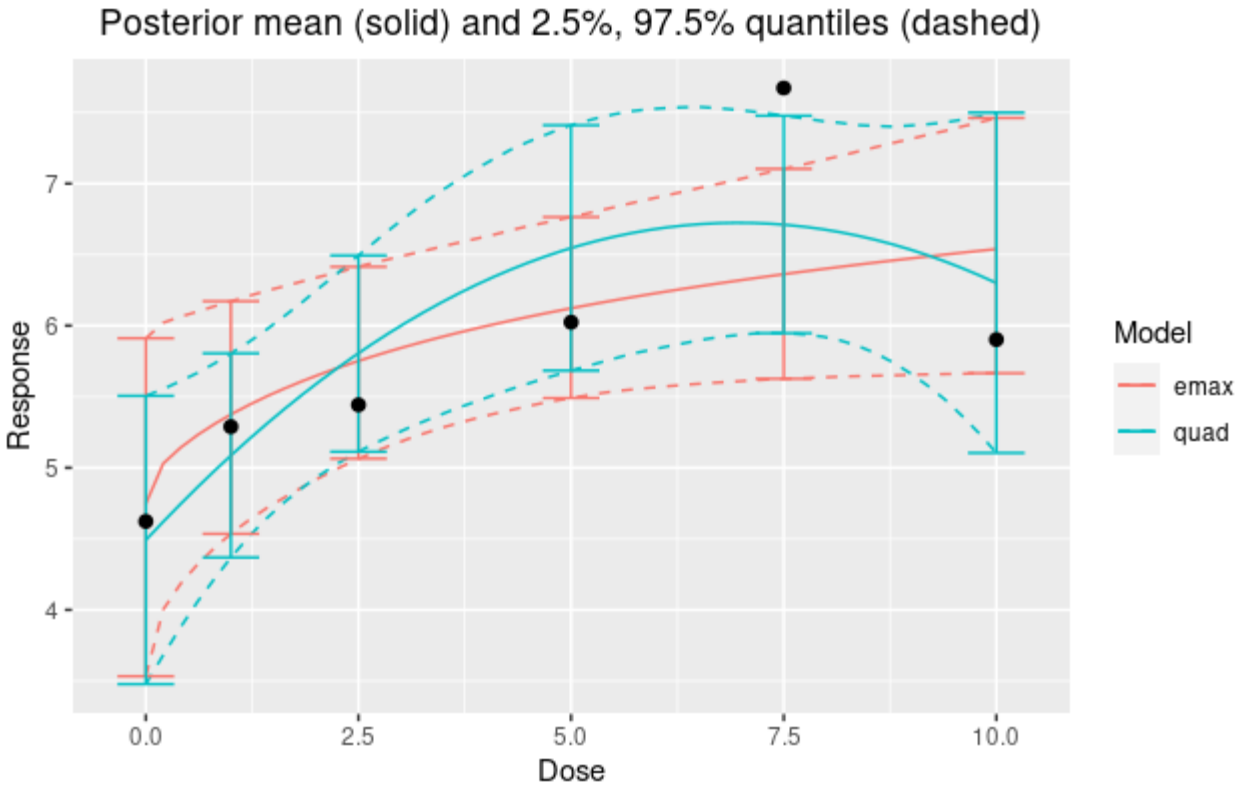
EMAX Fit



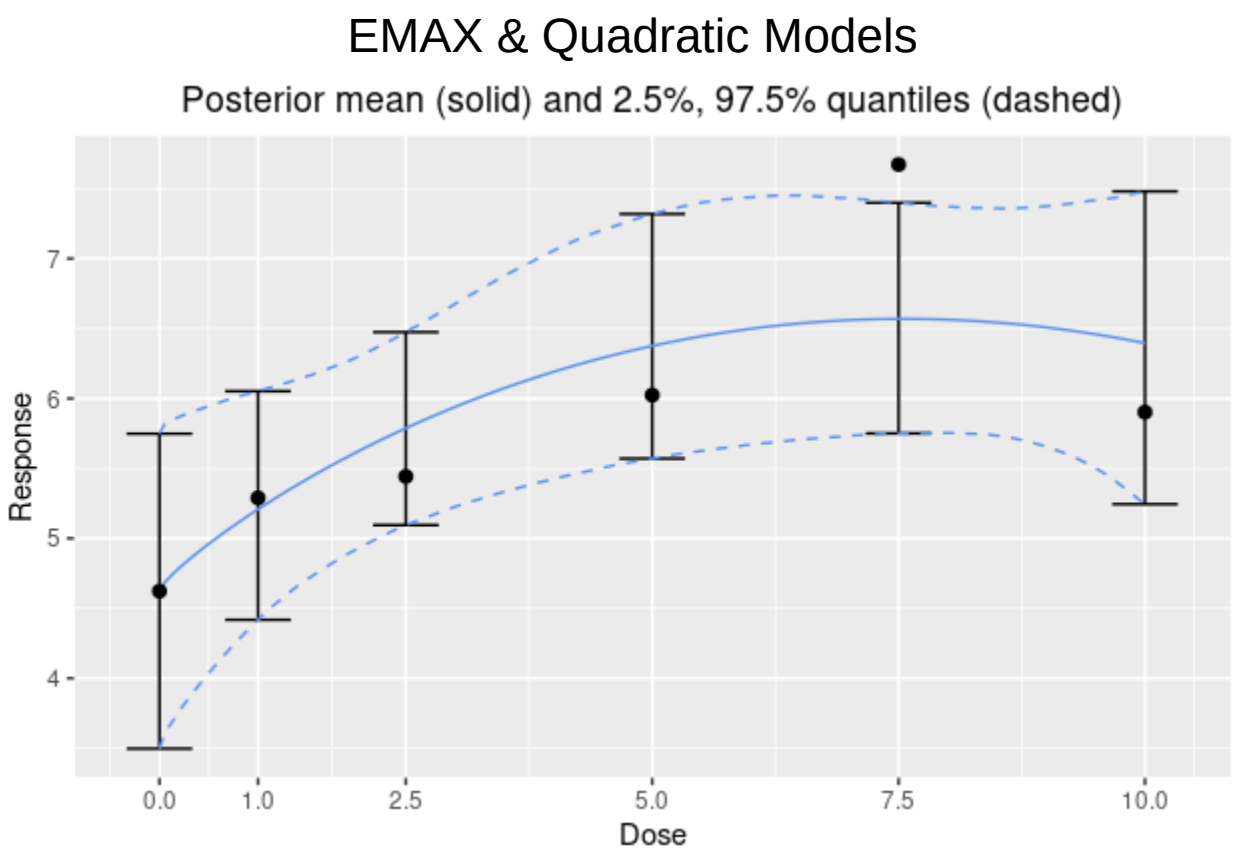
Quadratic Fit



Example: Dose Response Models



Example: Bayesian Model Averaging



Bayesian Model Averaging

Bayesian Model Averaging

- Bayesian model averaging is essentially a mixture prior over some quantity of interest.
- In our case, each component of the mixture prior is a different dose-response model.
  - Bayesian analog to the MCP-Mod
  - Bayesian version allows us to include informative priors and/or historical information into the analysis.

Bayesian Model Averaging

Let $\mu(d)$ represent the mean response at dose d.

We can construct a prior over M different parametric models for $\mu(d)$:

$$\pi\big(\mu(d)\big) = \sum_{m=1}^{M} \pi(\,\mu(d) \mid m\,)\,\pi(m)$$

Prior weight of model

Prior on the dose response for model m

Bayesian Model Averaging

The prior $\pi(\mu(d) \mid m)$ for each model is induced from a parametric model.

E.g. Linear Model: $\mu(d) = \beta_0 + \beta_1 d$

Obtain draws from $\pi(\beta_0, \beta_1)$ and insert into formula above.

To draw a sample from $\pi\big(\mu(d)\big)$ (full Bayesian model averaging prior):

1. Randomly select a model from $\pi(m)$
2. Randomly draw a set of parameters from that model's prior
3. Obtain $\mu(d)$ given the parameters drawn in 2.

Bayesian Model Averaging Posterior

$$p(\mu(d) \mid y) = \sum_{m=1}^{M} p(\mu(d) \mid y, m)\, p(m \mid y)$$

Marginal likelihood of data for model m (integrate over prior)

$$p(m \mid y) = \frac{p(y \mid m)\,\pi(m)}{\sum_{m^*} p(y \mid m^*)\,\pi(m^*)}$$

Bayesian Model Averaging

The **posterior** $\mathrm{p}(\mu(d) \mid m, \boldsymbol{y})$ for each model is induced from a parametric model.

E.g. Linear Model: $\mu(d) = \beta_0 + \beta_1 d$

Obtain draws from $\mathrm{p}(\beta_0, \beta_1 \mid \boldsymbol{y})$ and insert into formula above.

To draw a sample from $\mathrm{p}(\mu(d) \mid \boldsymbol{y})$ (Bayesian model averaging **posterior**):

1. Randomly select a model from $p(m \mid \boldsymbol{y})$
2. Randomly draw a set of parameters from that model's posterior
3. Obtain $\mu(d)$ given the parameters drawn in 2.

Difficulties of BMA

Calculating the marginals: $p(y \mid m)$

- Closed forms usually not available
- Monte Carlo estimates often have high variability and are therefore unreliable.
- One can avoid calculating marginals through one large MCMC chain, e.g., reversible jump MCMC (Green, 1995).
  - Difficult to ensure mixing
  - Computationally intense, often need custom MCMC samplers.

Sensitivity to diffuse priors

Philosophical Thoughts of BMA

- Do the classical weights using $p(y \mid m)$ really give us what we want?
- Essentially $p(y \mid m)$ is “how likely is it that this prior generated the parameters which generated the observed data”
- Don’t we really want the weights to reflect “which model fits the data best?”

An Alternative Weight

- Ando & Tsay (2010) replace $p(y \mid m)$ with $\exp(p^*(y \mid m))$ where $p^*(y \mid m)$ is an estimate of the posterior log-predictive likelihood of the observed data for model m.
- This is justified using a Kullback-Leibler argument (comparing the empirical and posterior predictive distributions).

Pros

- MCMC can be fit separately for each candidate model.
- Weights are less sensitive to diffuse/non-informative prior choices.
- Calculation of weights can be obtained directly from MCMC output.

Cons

- Breaks the canonical Bayes' Formula
  - (Is that a bad thing?)
- The estimate of $p^*(y \mid m)$ is biased
  - Corrections are suggested by Ando & Tsay (2010), assuming i.i.d data.

BMA and Dose-Response

- Gould (2019) proposed BMA-Mod which is the Bayesian analog of MCP-Mod.
- Applies Bayesian model averaging with the weights of Ando & Tsay (2010) to dose-response modeling.
- Includes a number of interesting examples.

Longitudinal Dose-Response Models

(MANUSCRIPT IN PROGRESS)

Lilly

Why Longitudinal models?

- Potentially improve decision making earlier
  - Use all available information (e.g. from not-yet-completers)
- Understand the longitudinal response profile for each dose.

Longitudinal Dose Response

We consider a class of longitudinal dose-response models of the form

$$\mu(d,t) = \alpha + g(d) \times f(t)$$

where $g(d)$ is a dose-response model, $f(t)$ is a longitudinal profile at time t.

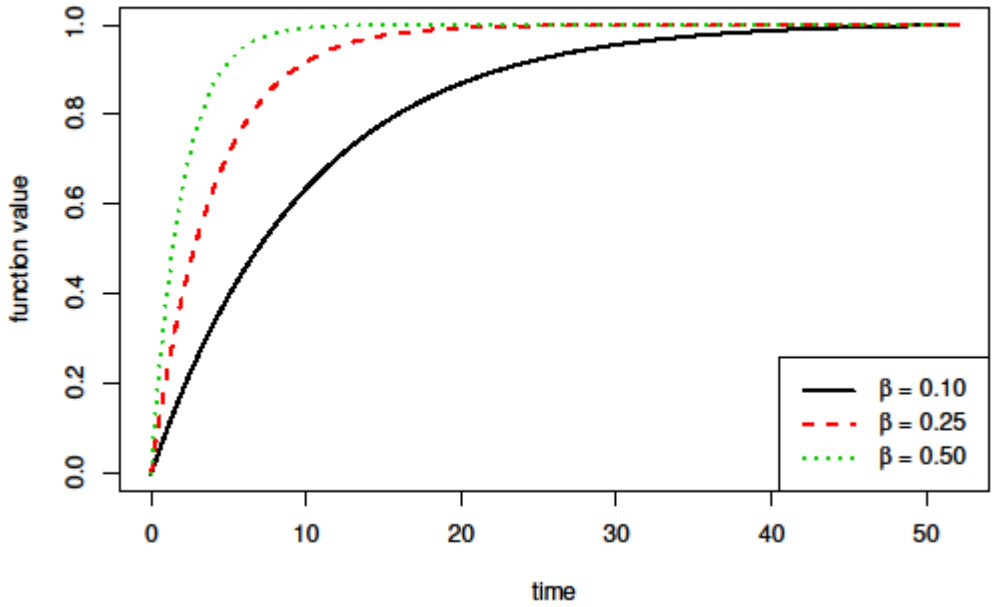
$f(t)$ must satisfy the following conditions:

- Continuous
- $f(0) = 0$
- $0 \leq f(t) \leq 1$
- $\max_t f(t) = 1$

ITP Model

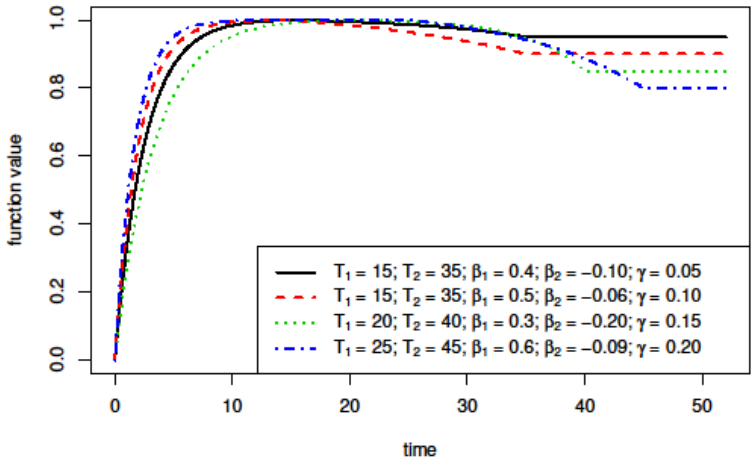
- Fu and Manner (2010)

$$f(t) = \frac{1 - \exp(-\beta t)}{1 - \exp(-\beta T)}$$



IDP Model (Pallavi Ray)

$$f(t) = f_1(t)I(0 \le t < T_1) + f_2(t)\, I(T_1 \le t < T_2) + f_2(T_2)I(T_2 \le t \le T)$$

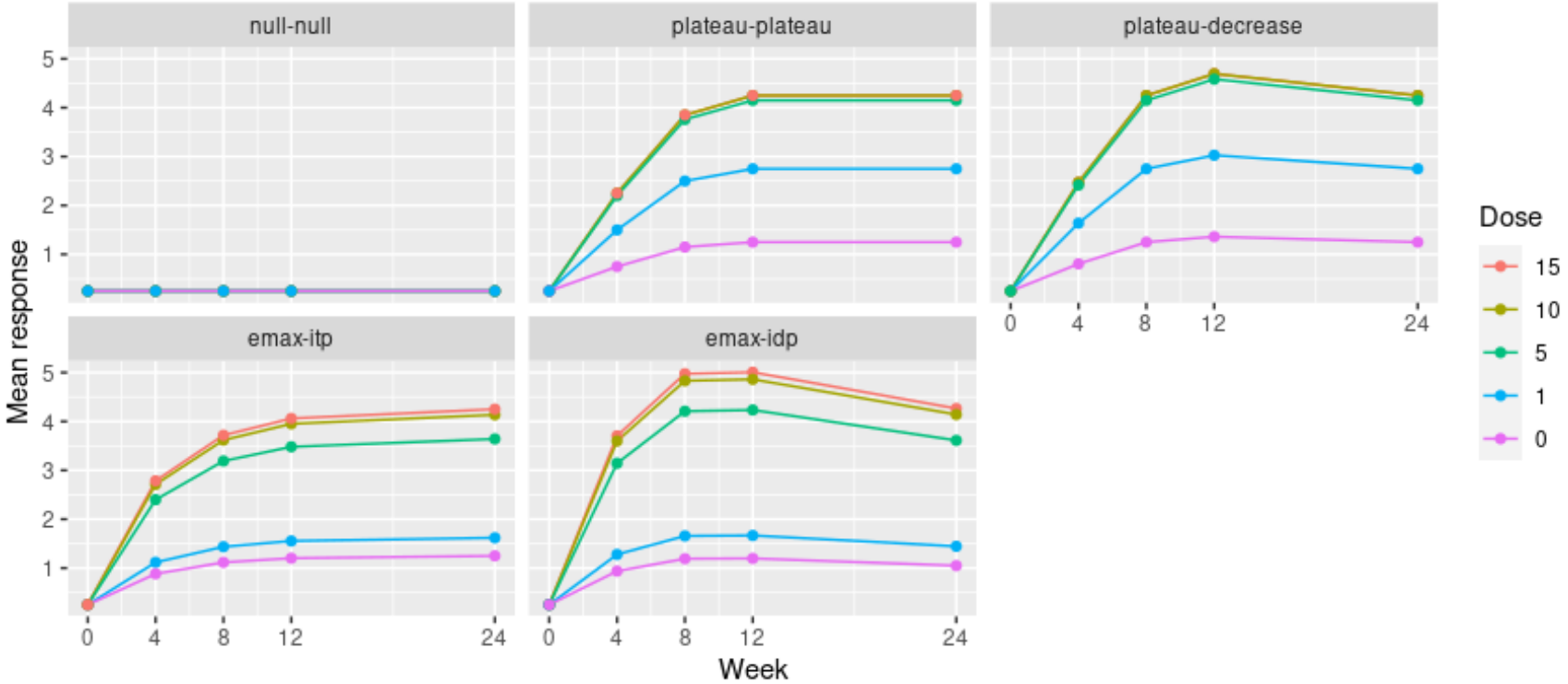


$$f_1(t) = \frac{1 - \exp(-\beta_1 t)}{1 - \exp(-\beta_1 T_1)}$$

$$f_2(t) = 1 - \gamma \frac{1 - \exp(-\beta_2(t - T_1))}{1 - \exp(-\beta_2(T_2 - T_1))}$$

Simulations

Trial Simulation Scenarios

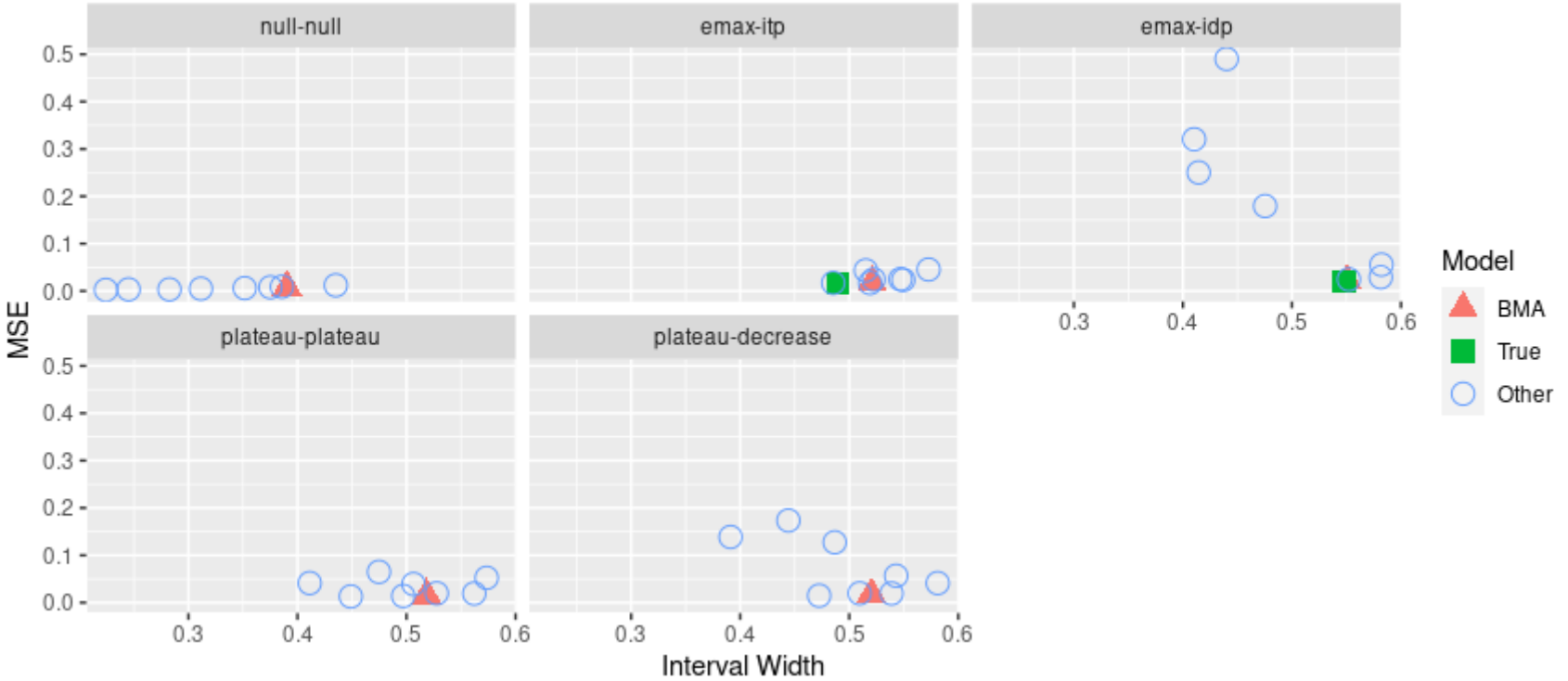


Setup

- 1:1:1:1:1 randomization over doses
- Interim for early futility or efficacy
  - Interim timing at 50% and 75% completers
- 25 and 50 subjects per arm
- Enrollment of 10 and 30 patients per week
- Analysis: Bayesian model averaging with dose response models: quadratic, log-quadratic, EMAX, exponential crossed with longitudinal ITP and IDP models.
  - 8 total models with equal prior weight.

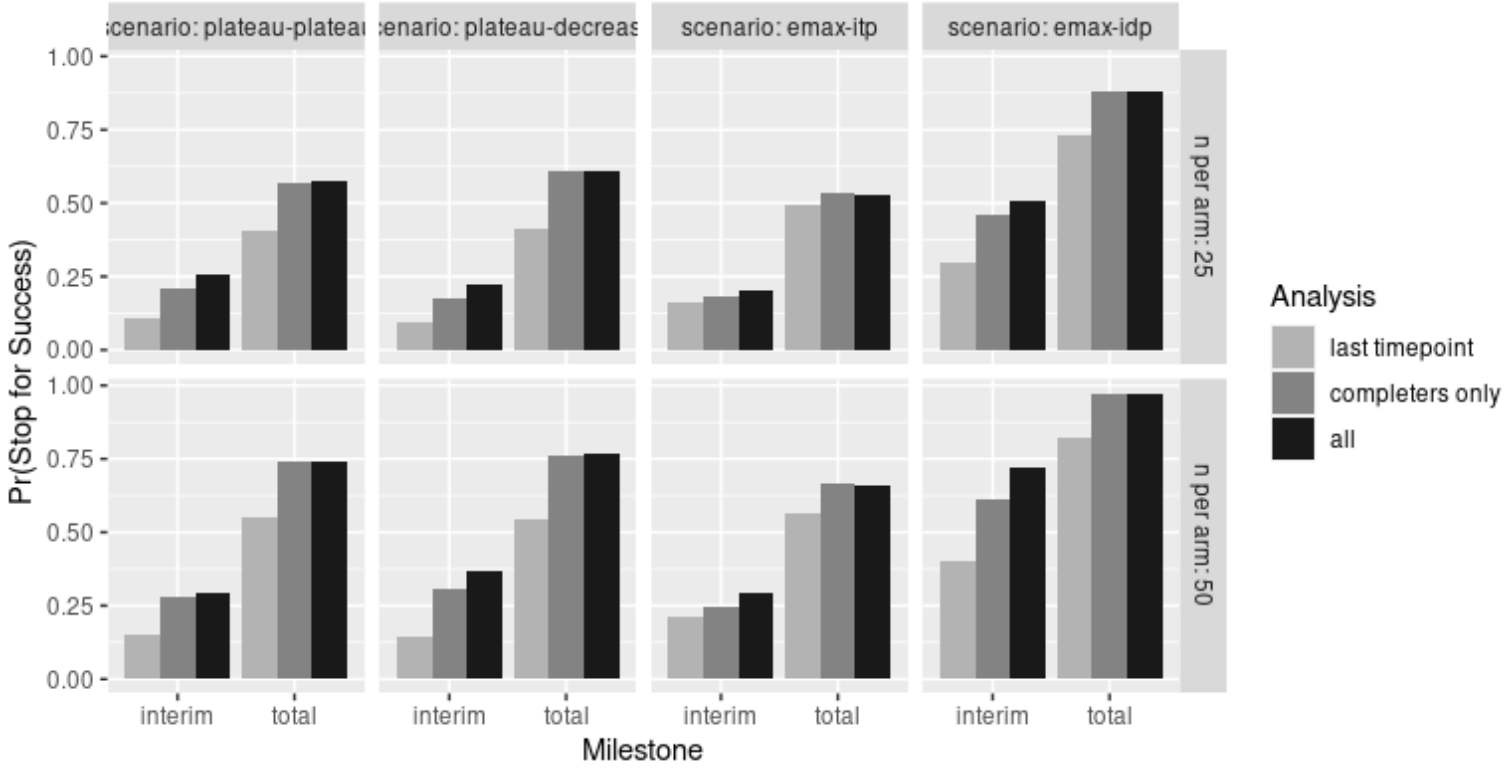
MSE

MSE vs Interval Width



Posterior Weights

Interim Stopping



Future Work

- More complex correlation structures for the longitudinal component.
- Comparison with non-parametric methods

dreamer R Package

dreamer R package

- https://github.com/rich-payne/dreamer

References

- Ando, T., & Tsay, R. (2010). Predictive likelihood for Bayesian model selection and averaging. *International Journal of Forecasting*, *26*(4), 744-763.
- Gould, A. L. (2019). BMA-Mod: A Bayesian model averaging strategy for determining dose-response relationships in the presence of model uncertainty. *Biometrical Journal*, *61*(5), 1141-1159.
- Fu, H., & Manner, D. (2010). Bayesian adaptive dose-finding studies with delayed responses. *Journal of biopharmaceutical statistics*, *20*(5), 1055-1070.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, *82*(4), 711-732.
- Sacks, L. V., Shamsuddin, H. H., Yasinskaya, Y. I., Bouri, K., Lanthier, M. L., & Sherman, R. E. (2014). Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000-2012. *Jama*, *311*(4), 378-384.