Bayesian Model Averaging for the Partial Ordering Continuous Reassessment Method

Luka Kovacevic

September 8, 2021

1 Methodology

Suppose we have M potential orderings of k prespecified dose levels, $d_1, d_2, \ldots, d_k$. We begin by assuming that the dose-toxicity relationship which corresponds to each combination of ordering, $m \in \{1, \ldots, M\}$, and dose level, $i \in \{1, \ldots, k\}$, can be modelled by the following single parameter working model for the probability of toxicity τ_i,

$$\tau_i = \psi_m(d_i, a) = \alpha_{mi}^{\exp(a)}, \tag{1}$$

where a is the model parameter which has the same distribution regardless of dose level and ordering and α_{mi} is the skeleton probability for this combination of model and dose. The skeleton probabilities are selected using the

The working model specified in equation (1) assumes that the dose-toxicity relationship is monotonic for a particular combination of ordering m and dose d_i. This is significant since it allows for multiple models to be specified each of which corresponds to a separate potential ordering, however, there are a number of ways that the various models can be handled to account for uncertainty in our knowledge of orderings.

Using these model orderings we can estimate the probability of toxicity for each dose level denoted as $\widehat{\tau}_i$. Hence, by obtaining

$$\arg\min_{d_i} |\widehat{\tau}_i - \theta|, \tag{2}$$

where θ is the TTR, the dose for the next cohort of patients in the trial is allocated. Following this method, we continue until our stopping rules are satisfied and the final recommended dose is the one that is considered to be the MTD for this trial.

1.1 Continuous Reassessment Method for Partial Ordering

Under the Bayesian framework of the CRM, we set a prior distribution for our set of parameters, $f(a)$, and a prior probability for each model $p(m)$. The POCRM uses model probabilities to choose the model that is most likely to be true and doses are allocated according to this selection which is outlined in equation (5). If some model probabilities are equal then the model used for the allocation is chosen randomly among those that have the maximum probability.

Since the outcome being observed is whether an individual experiences a DLT the outcome is binary, and hence, we obtain the following likelihood under model m after the inclusion of j patients in the trial,

$$L_m(a|\Omega_j) = \prod_{l=1}^{j} \psi_m^{y_l}(x_l, a) \left\{1 - \psi_m(x_l, a)\right\}^{1-y_l},$$

where Ω_j denotes the observed data set. Given the likelihood, the posterior density for the parameter a under each ordering is as follows,

$$f_m(a|\Omega_j) = \frac{L_m(a|\Omega_j) f(a)}{\int_{\mathcal{A}} L_m(a|\Omega_j) f(a) da}. \tag{3}$$

The posterior model probabilities for each model can also be obained as given by,

$$p(m|\Omega_j) = \frac{p(m)\int_{\mathcal{A}} L_m(a|\Omega_j)f(a)da}{\sum_m p(m)\int_{\mathcal{A}} L_m(a|\Omega_j)f(a)da}. \tag{4}$$

The model used to apply the Bayesian CRM is selected using,

$$b = \arg\max_m p(m|\Omega_j). \tag{5}$$

Then, the posterior density corresponding to the selected model used to estimate a posterior mean for the parameter a as follows,

$$\widehat{a}_b = \int_{\mathcal{A}} a f_b(a|\Omega_j)da, \tag{6}$$

which can then be plugged directly into the working model to obtain an estimate for the probability of toxicity for each dose,

$$\widehat{\tau}_i = \psi_b(d_i, \widehat{a}_b), \tag{7}$$

from which we can then obtain a recommendation for the next dose level using equation (2).

1.2 Bayesian Model Averaging

Suppose that rather than selecting a single model or ordering for each dose allocation in the trial we would prefer to take all the orderings into account before making the next decision on dose allocation. This is particularly attractive in cases where model probabilities are similar (e.g. 0.39 and 0.40) as choosing a single model in such cases would completely ignore the fact that alternative model is equally likely. Bayesian model averaging would allow for such an implementation by combining the available posterior information about the parameter a and the M model probabilities.

1.2.1 Point Estimate

If we are interested in the

The mean of the working model across the posterior distribution of the parameter a under each of the potential models is first computed by the following,

$$\widehat{\tau}_{mi} = \int_{\mathcal{A}} \psi_m(d_i, a) f_m(a|\Omega_j)da, \tag{8}$$

where $f_m(a|\Omega_j)$ is the posterior density of a, $\psi_m(d_i, a)$ is the working model and τ_{mi} is the probability of toxicity assuming ordering m. Using the posterior model probabilities, $p(m|\Omega_j)$, we can apply BMA to the estimates of the risk of toxicity which gives,

$$\widehat{\tau}_i = \sum_m p(m)\widehat{\tau}_{mi}. \tag{9}$$

The combined estimate, $\widehat{\tau}_i$, which is the estimate for the probability of toxicity under dose d_i regardless of model, is used to allocate the next dose for experimentation as shown in equation (2).

1.2.2 Mixture Distribution

BMA can also be applied to the posterior distribution of our set of parameters, a. This method makes it more straightforward to implement various stopping rules such as overdose controlling and allows for uncertainty in a to be taken into account in the final estimate for the risk of toxicity.

First, suppose that we are now interested in obtaining the density function for a particular probability of toxicity, τ_i for a specific dose level, i. Then, from the working model we have the following,

$$\begin{aligned} &\tau_i = \psi_m(d_i, a) = \alpha_{mi}^{\exp(a)}, \\ &\Rightarrow \log\left\{\frac{\log(\tau_i)}{\log(\alpha_{mi})}\right\} = a = \psi_m^{-1}(d_i, \tau_i), \end{aligned} \tag{10}$$

Hence, since τ_i is a function of a we can obtain a density function for τ_i as follows,

$$\begin{aligned} f_m(\tau_i|\Omega_j) &= \left| \frac{d}{d\tau_i}(\psi_m^{-1}(d_i, \tau_i)) \right| f_m(\psi_m^{-1}(d_i, \tau_i)|\Omega_j), \\ &= \left| \frac{1}{\tau_i \log(\tau_i)} \right| f_m(\psi_m^{-1}(d_i, \tau_i)|\Omega_j). \end{aligned} \tag{11}$$

BMA is then applied to the set of posterior distributions of the probability of toxicity at each dose to obtain the following,

$$g(\tau_i|\Omega_j) = \sum_m p(m|\Omega_j) f_m(\tau_i|\Omega_j), \tag{12}$$

which is the combined posterior distribution of the probability of toxicity for a particular dose level d_i and is a function of a particular probability of toxicity in which we are interested in (e.g. TTR).

There are two available approaches to dose allocation in this setting. For the first approach we use the following,

$$\arg\max_{d_i} g(\tau_i = \theta|\Omega_j), \tag{13}$$

where θ is the TTR, and we select the dose with the maximum posterior density for further experimentation. Alternatively, we can compute the expected posterior probability of toxicity given by,

$$\widehat{\tau}_i = \int_0^1 \tau_i g(\tau_i|\Omega_j) d\tau_i, \tag{14}$$

which can then be used in conjunction with equation (2) to select the next dose for experimentation.

1.2.3 Overdose Controlling

Overdose controlling can be implemented to prevent allocation to overly toxic doses or doses that have a high risk of being overly toxic. This type of safety constraint can be expressed as follows,

$$\Pr[\psi(d_i, a) > \theta] < c_{\text{od}}, \tag{15}$$

where θ is the target toxicity level and c_{od} is the specified constant for overdose control. If inequality (15) is violated, the i-th dose is deemed unsafe. In this implementation, only safe doses (i.e. only doses that satisfy this constraint) can be selected as the next dose. In the case where no doses are considered to be safe, the trial is stopped.

For the POCRM and point estimate approaches we first determine

$$a^* = \log\left\{ \frac{\log(\theta)}{\log(\alpha_{mi})} \right\} \tag{16}$$

which we can then substitute to obtain the following probability for the POCRM approach,

$$\Pr[\psi_m(d_i, a) > \theta] = \int_{-\infty}^{a^*} f_m(a|\Omega_j) da. \tag{17}$$

We can then apply BMA to obtain the following probability which is used to check the safety constraint in the case of the point estimate method,

$$\Pr[\psi(d_i, a) > \theta] = \sum_m p(m|\Omega_j) \Pr[\psi_m(d_i, a) > \theta] \tag{18}$$

Similarly, for the approaches utilising mixture distributions, we can substitute the target toxicity rate directly as follows,

$$\Pr[\psi(d_i, a) > \theta] = \int_\theta^1 g(\tau_i|\Omega_j), \tag{19}$$

which can directly be used for the safety constraint.

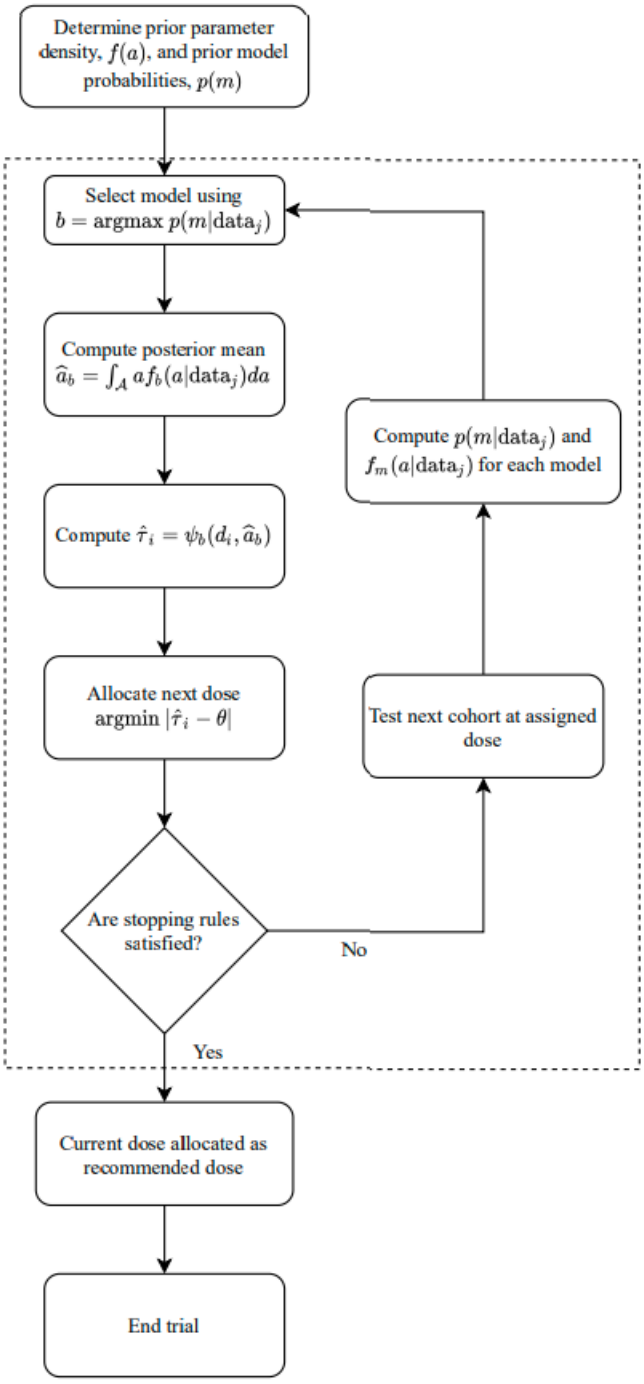

Figure 1: Flowchart outlining the continual reassessment method for partial ordering.

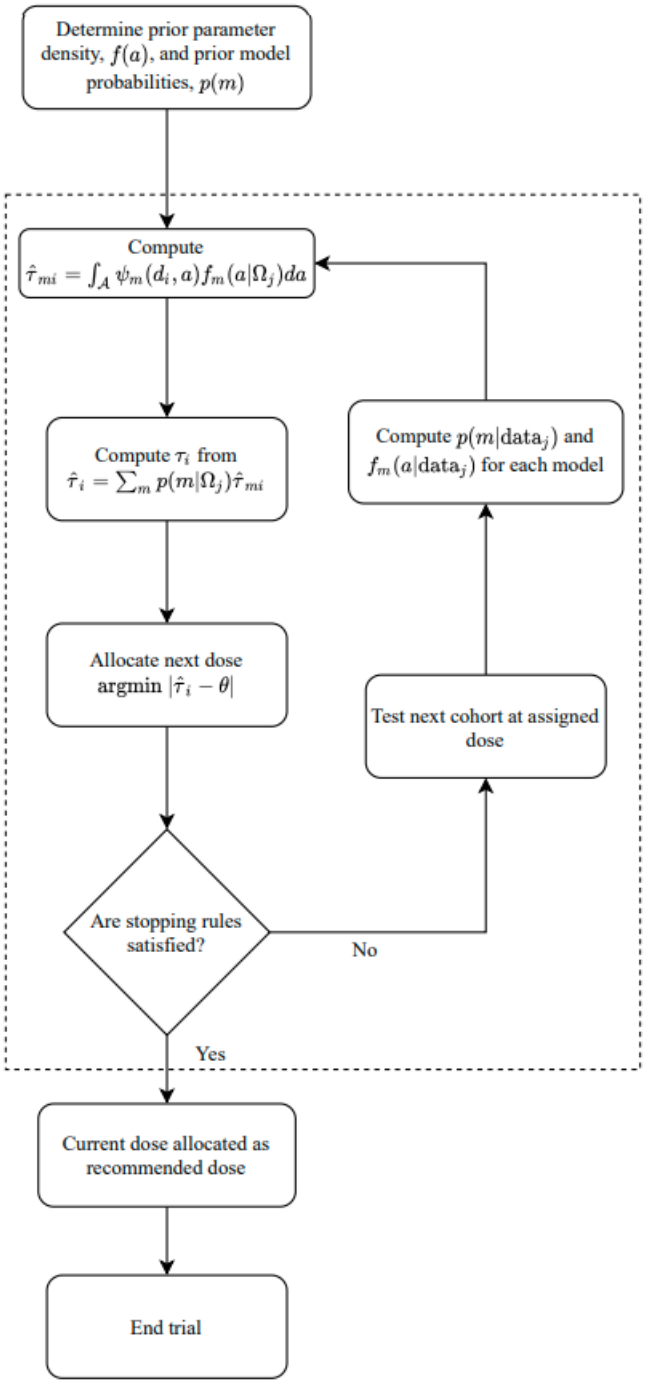

Figure 2: Flowchart outlining the point estimate implementation of BMA for the continual reassessment method for partial ordering.

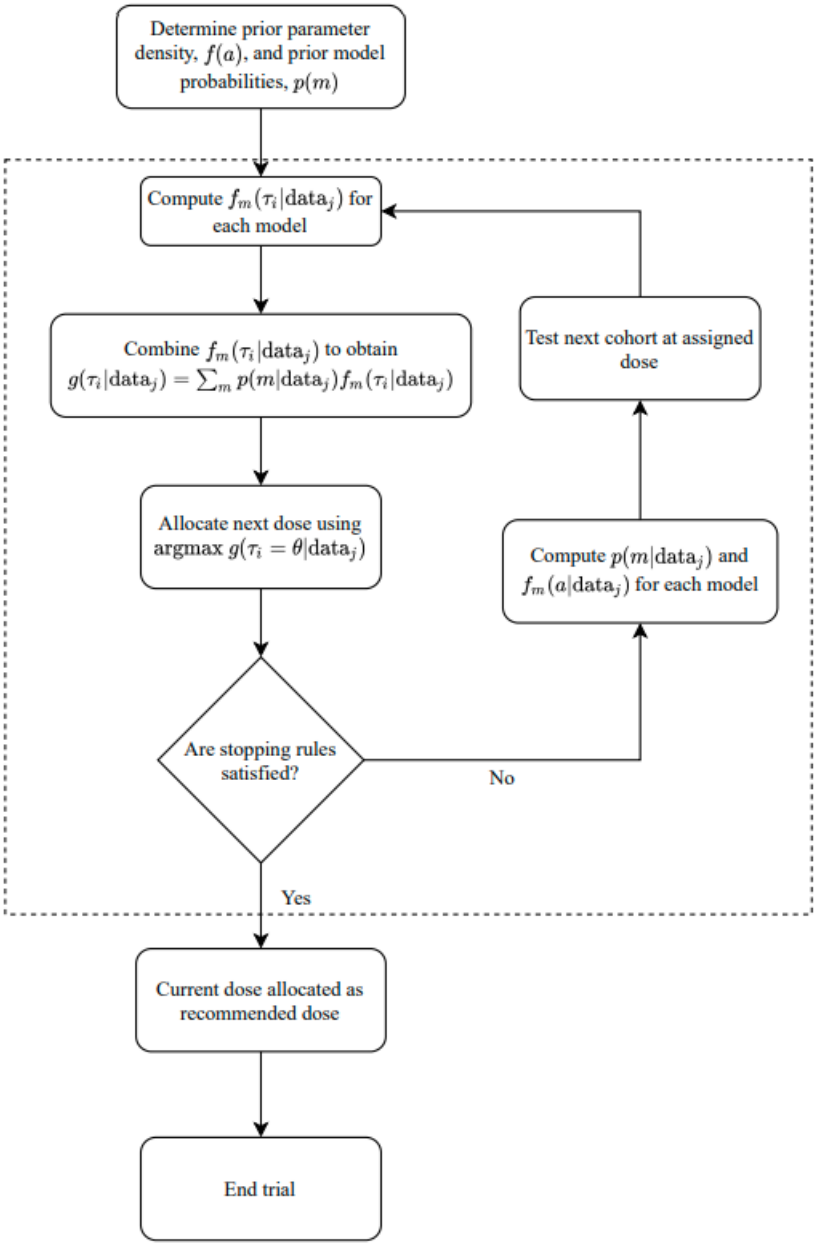

Figure 3: Flowchart outlining the mixture distribution implementation of BMA for the continual reassessment method for partial ordering which allocates doses via densities.

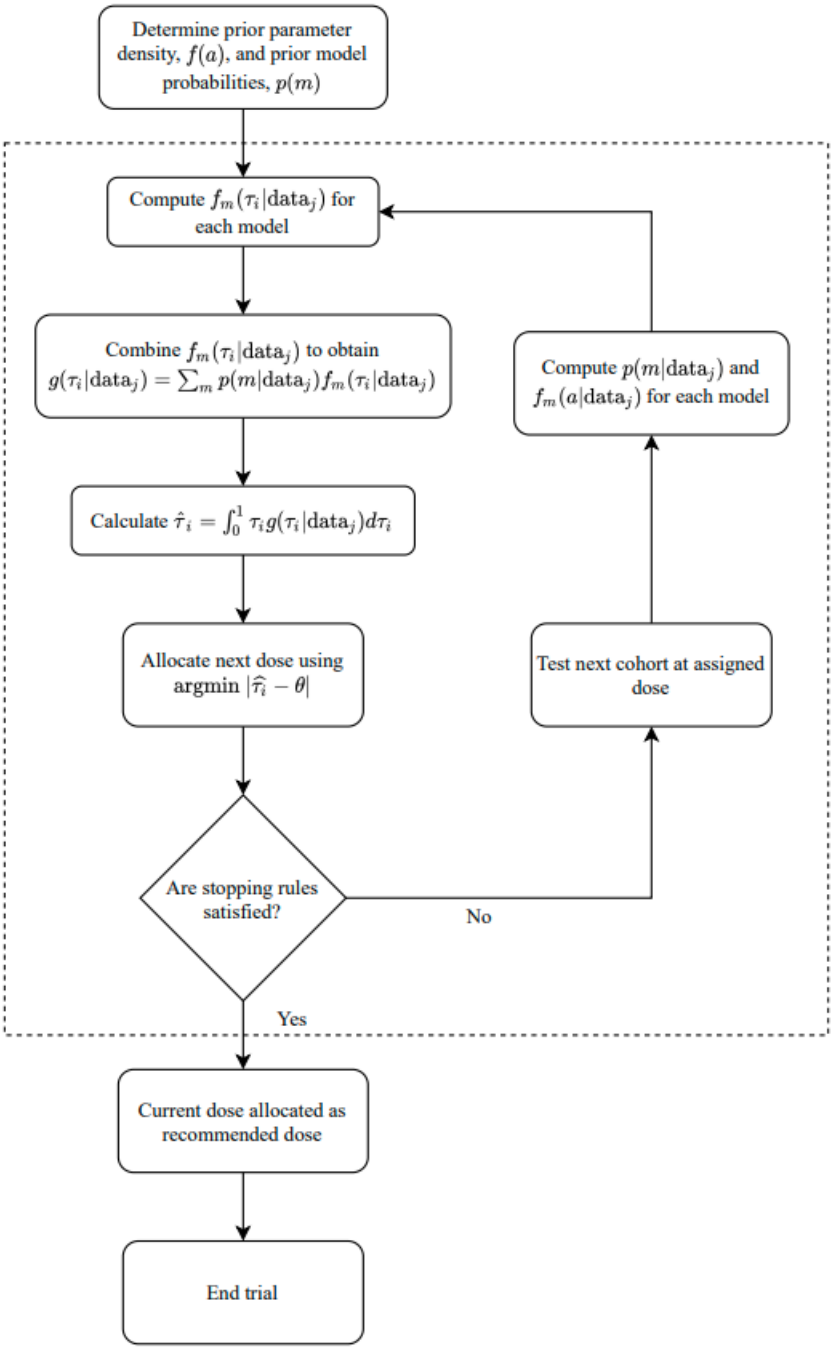

Figure 4: Flowchart outlining the mixture distribution implementation of BMA for the continual reassessment method for partial ordering which allocates doses via estimates of the probability of toxicities.