Using Subject Level Covariate Information in Bayesian Mixture Models for Basket Trials

Sneha Govande* Elizabeth H. Slate†

Abstract

A basket trial evaluates one or more treatments for efficacy among more than one cancer type in a single clinical trial. Though the treatment targets the common genetic aberration that causes different cancer types, the possible heterogeneity in the treatment effects poses challenge in modeling. Compared to traditional designs, basket trials can reduce the time required for testing and, by pooling across cancer types, they also allow the drugs to be tested for rare cancers. Basket trials are gaining increasing importance with advancements in precision medicine. Using covariate information has shown merit for improving efficacy in classification of the baskets. We incorporate subject-level biomarker information to aid identification of responsive and non-responsive baskets. We model subjects' responses using a two-component Bayesian mixture model where the mixture weights depend on a measure of similarly among subjects' biomarker values. We demonstrate the performance of this model using simulation.

Key Words: Basket trials, Bayesian mixture model, Biomarker data, Covariate information, classification

1. Introduction

Different cancer types have commonly been identified based on tumor locations (histologies). However, with increasing knowledge about cancer biology, it has been possible to understand that different cancer types can be caused by a common gene aberration which can be addressed using targeted therapies and immunotherapies. Traditional phase 2 trials evaluate a treatment independently for each cancer type. Basket trials allow us to test the treatment for related tumor types in a single trial. The term basket trial broadly refers to the set-up where one or more treatments are tested for efficacy among multiple cancer types (baskets) in a single clinical trial. These trials offer several advantages compared to the traditional set up, and thus they are gaining increasing importance with advancements in precision medicine. Combining different cancer types together in one trial may prove to be more efficient, as it uses the common gene aberration to strengthen inference for each type, and thus allows study of rare cancer types while maintaining desirable statistical properties. Heterogeneity of treatment effects, however, may pose a challenge, as seen in previous trials like the Vemurafenib trial (Hyman et al., 2015).

2. Literature Review

A variety of methods for design and analysis of basket trials have been proposed. A traditional independent analysis applies a design of choice independently within each tumor type. This approach does not allow any information to be shared across the tumor types and hence leads to estimates with no shrinkage and higher sample size requirements compared to other approaches that allow information sharing. At the opposite extreme is pooling all the tumor types together and analyzing them as one group. Although capitalizing on the common etiology, this approach fails to recognize heterogeneity in the treatment effects

*Department of Statistics, Florida State University
†Department of Statistics, Florida State University

among the tumor types. The hierarchical Bayesian model suggested by Thall et al. (2003) governs the information borrowing across tumors through type-specific exchangeable treatment effects. The basket-specific estimates that result are shrunk toward the pooled mean, which efficiently incorporates an anticipated common behavior, but can mask heterogeneity, making it difficult to detect a responsive basket if there is only one. Many designs that use the hierarchical modeling approach have been proposed, including the Bayesian adaptive design with frequent interim analysis of Berry et al. (2013), the calibrated hierarchical Bayesian model approach by Chu and Yuan (2018a), which models a variance parameter to control the shrinkage, and the exchangeability-nonexchangeability (EXNEX) design by Neuenschwander et al. (2016) that uses a robust mixture extension of an exchangeability model allowing each stratum specific parameter to be exchangeable with other similar strata parameters or nonexchangeable with all of them. Chu and Yuan (2018b) propose using joint modeling of treatment response and longitudinal biomarker covariate data to control information borrowing in their Bayesian latent subgroup design for basket trials (BLAST). Zhou and Ji (2020) suggest the Robust Bayesian Hypothesis Testing (RoBoT) method that also assumes a latent subgroup structure. Ventz et al. (2017) propose a general class of Bayesian response adaptive designs for multi-arm trials with biomarker defined subgroups and multiple malignancies using hierarchical models. Hobbs and Landin (2018) propose a novel methodology for sequential basket trial design formulated with Bayesian monitoring rules based on a novel hierarchical modeling strategy for sharing information among a collection of discrete potentially nonexchangeable subtypes. Liu et al. (2017) propose a two stage design where homogeneity is assessed at the first stage and information is borrowed accordingly in the second stage.

We present more detailed descriptions of selected approaches below. Let B be the number of tumor types (baskets), with n_b subjects in the b^{th} basket. Let $y_b = (y_{1b}, y_{2b}, \ldots, y_{n_b b})$ be the vector of subject-level responses and α_b be the basket-specific parameter associated with the distribution of y_b in basket b, where $b = 1, 2, \ldots, B$. Also let $f(\cdot)$ denote a pdf associated with the random variable indicated by its argument.

Independent Approach: A clinical trial design is applied independently in each of the B baskets so that the parameters $\{\alpha_b\}$ are modeled independently; see (1). The independent modeling approach was used in the Vemurafenib trial by Hyman et al. (2015) with Simon's two stage design (Simon, 1989) applied independently in every basket. Under the Bayesian framework, the basket specific parameters are assigned independent prior distributions and are estimated using the posterior distributions given observed response data.

$$\begin{aligned} y_{ib} \mid \alpha_b &\sim f(y_{ib}|\alpha_b) \quad \text{iid}, i = 1, 2, \ldots, n_b, \\ \alpha_b &\sim f_b(\alpha_b) \quad b = 1, 2, \ldots, B. \end{aligned} \tag{1}$$

This approach does not take advantage of the commonality among the baskets, which means no information is borrowed and estimates $\hat{\alpha}_b$ are not subject to shrinkage. A go/no-go decision is made independently for each basket.

Bayesian Hierarchical Model (BHM): A common formulation of the BHM is shown in (2). The individuals' responses within basket b are again modeled as independent conditional on the basket-specific parameter α_b, but the α_bs are assumed to arise from a common distribution, typically parameterized with mean μ and precision τ, which are further assigned a prior distribution. Inference for μ, τ and $\{\alpha_b\}$ proceeds from the posterior

distribution.

$$\begin{aligned} y_{ib} \mid \alpha_b &\sim f(y_{ib}|\alpha_b), \text{ iid}, i = 1, 2, \ldots, n_b, \\ \alpha_b \mid \mu, \tau &\sim f(\alpha_b|\mu, \tau), \text{ iid}, b = 1, 2, \ldots, B, \\ (\mu, \tau) &\sim f(\psi). \end{aligned} \tag{2}$$

The model assumes exchangeability of basket specific parameters. This assumption may fail to hold when the treatment effects are heterogeneous. The amount of borrowing and thus the degree of shrinkage is controlled by the precision parameter of the distribution of basket specific parameters. Therefore, this parameter is also called a shrinkage parameter. The greater the precision (i.e., smaller the variance), the higher the degree of shrinkage.

Bayesian Hierarchical Model with Covariates: This modification of the BHM in (2) incorporates a basket-level covariate, x_b, $b = 1, 2, \ldots, B$, and the response parameters α_b are modeled as a function of x_b. An example formulation is given in (3), where the covariate x_b influences the basket parameter α_b through the link function $g(\cdot)$. The basket-level errors $\{\delta_b\}$ would commonly be modeled as having mean zero and, in the Bayesian setting, a prior distribution is specified for the regression parameters β_0 and β_1.

$$\begin{aligned} y_{ib} \mid \alpha_b &\sim f(y_{ib}|\alpha_b), \text{ iid}, i = 1, 2, \ldots, n_b, \\ \theta_b &= g(\alpha_b) = \beta_0 + \beta_1 x_b + \delta_b \\ \delta_b &\sim f(\delta_b), \text{ iid}, b = 1, 2, \ldots B \\ (\beta_0, \beta_1) &\sim f(\psi). \end{aligned} \tag{3}$$

Product Partition Model: A product partition model (PPM) (Barry and Hartigan, 1992) has been used to model data with nonexchangeability. Barry and Hartigan (1992) proposed the PPM for change point problems where the sequence of observations observed at consecutive points in time is partitioned into clusters. A different probability model is assumed to hold within each of the clusters. Part of the inference problem is discovery of the partition. Let $\rho = (S_1, S_2, \ldots, S_{k_n})$ denote a partition of n experimental units into k_n subsets, $y = (y_1, y_2, \ldots, y_n)$ denote the response vector with y_j denoting the response for unit j, $y^k = (y_j, j \in S_k)$ denote the response data arranged by clusters, $e = (e_1, e_2, \ldots, e_n)$ denote the cluster membership indicator with $e_j = k$ if $j \in S_k$, where $k = 1, 2, \ldots, k_n$ and $j = 1, 2, \ldots, n$. The values of $(k_n, e_1, \ldots, e_n)$ describe a partition up to permutation of the cluster labels. The number of clusters k_n is unknown. A prior probability model on the partition, $f(\rho)$, implies a prior on the number of clusters, k_n, in the partition. The PPM constructs $f(\rho)$ by introducing a cohesion function $c(A) \geq 0$ for $A \subseteq \{1, 2, \ldots, n\}$ that measures how tightly grouped the elements in A are thought to be. The choice of $c(A)$ depends on the prior belief about the partitions. The PPM then has the form

$$f(\rho) \propto \prod_{k=1}^{k_n} c(S_k) \quad \text{and} \quad f(y, \eta|\rho) = \prod_{k=1}^{k_n} f_k(y^k \mid \eta_k) f(\eta_k), \tag{4}$$

where $f_k(\cdot \mid \eta_k)$ is the model specific for cluster k and depends on parameters η_k. All inference concerning ρ is made from the posterior distribution $f(\rho \mid y)$ (Dahl, 2009).

3. Bayesian Partition Model with Covariates

The goal for our model is to account for possible non-exchangeability when analyzing the data from basket trials. We seek to incorporate subject-level covariate information that informs classification of the baskets into latent clusters that aid basket-wise go/no-go decisions.

As earlier, let $y_b = (y_{1b}, y_{2b}, \ldots, y_{n_b b})$ be the vector of responses for the n_b individuals in basket b, and, additionally, let $x_b = (x_{1b}, x_{2b}, \ldots, x_{n_b b})$ be the corresponding covariate values for these individuals, $b = 1, 2, \ldots, B$. The covariate could be a determination of the subject's prognosis or a measurement of a biological indicator of the subject's propensity for response to the treatment, for example. Here we assume that the covariate value is informative about treatment efficacy and will exhibit a cluster structure similar to that of the response. Let $y = (y_1, y_2, \ldots, y_B)$ and $x = (x_1, x_2, \ldots, x_B)$ be the complete vectors of response values and covariate values, respectively, for all subjects. To account for possible heterogeneity in treatment effects, we assume that each of the B baskets comes from a mixture of two distributions corresponding to two latent clusters of, respectively, responsive baskets and non-responsive baskets. A goal is to classify the B baskets into two clusters (one of which may be empty) such that baskets that have similar subject-level covariate values will be in the same cluster. Let e_b denote the cluster membership indicator for basket b, with $e_b = j$ if the cancer type b belongs to cluster j, where $j = 1, 2$. Let $e = (e_1, e_2, \ldots, e_B)$ be the vector of cluster membership indicators, and let ρ denote a partition of B baskets into two clusters. The vector e of cluster membership indicators can be used to identify a corresponding partition (a partition ρ corresponds to two cluster membership vectors that differ only in the labeling of the clusters).

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_B)$ be the basket-specific parameters, $\theta = (\theta_1, \theta_2)$ be cluster-specific parameters and define $\psi = (\alpha, \theta)$. Conditional on the covariates x, we model the distribution of subject responses y as

$$\begin{aligned} f(y \mid \psi, x) &= \sum_{\rho} f(y, \rho \mid \psi, x) \\ &= \sum_{\rho} f(y \mid \psi, \rho) f(\rho \mid x) \\ &= \sum_{\rho} \left\{ \prod_{b=1}^{B} f(y_b \mid \alpha_b, \rho, \theta) \right\} \times f(\rho \mid x). \end{aligned} \tag{5}$$

The second equality in (5) indicates that the distribution of the response y depends on the covariate values only through the probability model for partitions. We propose to model the effect of x on the partition of the baskets ρ through a non-negative similarity function in the form of an auxiliary probability density function for x as suggested by Müller et al. (2011). Thus we have $f(\rho \mid x) \propto f(x \mid \rho) f(\rho)$, where we introduce $f(x \mid \rho)$ to quantify the combined similarity among x within the components of the partition ρ and $f(\rho)$ is the prior distribution on partitions in (4) that is uninformed by x. We require that this auxiliary distribution $f(x \mid \rho)$ incorporate within-cluster and within-basket dependence.

Let x^j denote the covariate values from cluster $j, j = 1, 2$, i.e., the covariate values

from the subjects in baskets in cluster j. Then our model for the partition becomes

$$\begin{aligned}
f(\rho \mid x) &\propto f(x \mid \rho) f(\rho) \\
&= f(x^1 \mid \rho) f(x^2 \mid \rho) f(\rho) \\
&= \prod_{j=1}^{2} \int f(x^j \mid \rho, \eta_j) f(\eta_j) d\eta_j f(\rho) \\
&= \prod_{j=1}^{2} \int \Big[\prod_{\{b:e_b=j\}} f(x_b|\eta_j) \Big] f(\eta_j) d\eta_j f(\rho) \\
&= \prod_{j=1}^{2} \int \Big[\prod_{\{b:e_b=j\}} \int f(x_b|\xi_b, \eta_j) f(\xi_b|\eta_j) d\xi_b \Big] f(\rho) \times f(\eta_j) d\eta_j f(\rho) \\
&= \prod_{j=1}^{2} \int \Big[\prod_{\{b|e_b=j\}} \int \Big\{ \prod_{i=1}^{n_b} f(x_{ib}|\xi_b, \eta_j) \Big\} \times f(\xi_b|\eta_j) d\xi_b \Big] f(\eta_j) d\eta_j f(\rho). \qquad (6)
\end{aligned}$$

Here $\{\eta_j\}$ are cluster-specific parameters and $\{\xi_b\}$ are basket-specific parameters that govern the distribution of x. Equations (5) and (6) together specify the proposed model that we call the Bayesian Partition Model with Covariates (BPMx).

The particular formulation of the BPMx that we consider is for binary response y_{ib} and continuous covariate x_{ib}. As shown in (7), we model y_{ib} as arising from a Bernoulli distribution with basket-specific mean α_b, and α_b as arising from a beta distribution governed by the cluster assignment of basket b. The covariate values inform the distribution of the partition of the baskets into two clusters (responsive and non-responsive). In the absence of the covariate information, we assume all partitions are equally likely, i.e., $f(\rho)$ in (6) is constant.

$$\begin{gathered}
y_{ib} \mid \alpha_b, e_b = j \;\sim\; \text{Bernoulli}(y_{ib} \mid \alpha_b), \text{ iid, } i = 1, 2, \ldots, n_b \\
\alpha_b \mid e_b = j, \theta_j = (\mu_j, V_j) \;\sim\; \text{Beta}(\alpha_b \mid \text{mean} = \mu_j, V_j) \\
f(\rho \mid x) \propto \prod_{j=1}^{2} \int \Big[\prod_{\{b|e_b=j\}} \int \Big\{ \prod_{i=1}^{n_b} \phi(x_{ib} \mid \xi_b, \sigma_x^2) \Big\} \times \phi(\xi_b \mid \eta_j, \sigma_\xi^2) d\xi_b \Big] \phi(\eta_j \mid \mu_\eta, \sigma_\eta^2) d\eta_j. \qquad (7)
\end{gathered}$$

Here V_j is a precision parameter of the beta distribution given by the sum of the usual shape parameters, and $\phi(x \mid \mu, \sigma^2)$ denotes the probability density function for the Gaussian distribution with mean μ and variance σ^2.

The Bayesian formulation of the BPMx in (7) is completed with a prior distribution for $(\mu_1, \mu_2, \sigma_x^2, \sigma_\xi^2, \sigma_\eta^2)$; the precisions V_1 and V_2 are specified for identifiability. Each beta mean μ_j arises from an equal mixture of two beta distributions, one corresponding to low and the other to high response probability. The inverse of the variance parameters are taken to be uniformly distributed.

4. Inference and Computations

Examining (5) and (6) reveals a significant advantage of incorporating the information from x about the partition in the form of an auxiliary probability distribution, $f(x \mid \rho)$. The model for the response y with the covariate-dependent prior for the partition is equivalent, for computational purposes, to a model for the augmented response (y, x) using the

auxiliary model for x and a prior for the partition that no longer depends on the covariate. Removing the dependence on x from the probability distribution for the partition aids computation substantially. Furthermore, we take $f(\rho) \propto 1$.

Because of the normality assumption in (7), $f(x \mid \rho)$ is a multivariate normal distribution of dimension $N = \sum_b n_b$. This auxiliary distribution can be expressed using random effects as follows:

$$\begin{aligned} x_{ib} &= \mu_\eta + \delta_j + \epsilon_b + \gamma_i \\ \delta_j &\sim \text{Normal}(0, \sigma_\eta^2), j = 1, 2 \\ \epsilon_b &\sim \text{Normal}(0, \sigma_\xi^2), b = 1, 2, \ldots, B \\ \gamma_i &\sim \text{Normal}(0, \sigma_x^2), i = 1, 2, \ldots, n_b \end{aligned} \tag{8}$$

where γ_i is the subject effect, ϵ_b is the basket effect and δ_j is the cluster effect. It follows from (8) that $\text{E}(x_{ib}) = \mu_\eta$ and $\text{var}(x_{ib}) = \sigma_\eta^2 + \sigma_\xi^2 + \sigma_x^2$. Moreover, the covariance structure for x is given by

$$\Sigma_x[i,j] = \begin{cases} \sigma_\eta^2 + \sigma_\xi^2 + \sigma_x^2 & \text{if } i = j; \\ \sigma_\eta^2 + \sigma_\xi^2 & \text{if subjects } i \text{ and } j \text{ are in the same basket;} \\ \sigma_\eta^2 & \text{if subjects } i \text{ and } j \text{ are from different baskets from the same cluster;} \\ 0 & \text{if subjects } i \text{ and } j \text{ are from different clusters.} \end{cases}$$

Inference is performed using Markov chain Monte Carlo; we use `Stan` (Stan Development Team, 2021) in our simulations in Section 5 and assess convergence using the Gelman-Rubin diagnostic (Gelman and Rubin, 1992). The basket-specific mean responses $\alpha_b, b = 1, 2, \ldots, B$, cluster-specific mean responses $\mu_j, j = 1, 2$ and variance parameters for the auxiliary covariate model σ_x^2, σ_ξ^2 and σ_η^2 can be estimated using summary of posterior samples.

Recall that a partition ρ corresponds to two cluster membership vectors that differ only in the labeling of the clusters. For example, in case of $B = 4$, $e = (1, 1, 1, 2)$ and $e = (2, 2, 2, 1)$ indicate the same partition for our purpose, with baskets 1, 2 and 3 in one cluster and basket 4 in the other cluster. To avoid this duplication of partitions, we fix the cluster assignment for the first basket to cluster 1 without loss of generality and classify the other baskets accordingly. This gives us 2^{B-1} cluster membership vectors, each identifying a unique partition. We estimate the cluster partition using the maximum *a posteriori* (MAP) value, which can be used to predict the cluster membership and aid the go/no-go decision. We also estimate the cluster membership probabilities $P_{bj} = \Pr(e_b = j \mid y, x)$ that a basket b comes from cluster j as follows:

$$\hat{P}_{bj} = \sum_{\rho : e_b = j} \Pr(\rho \mid x, y). \tag{9}$$

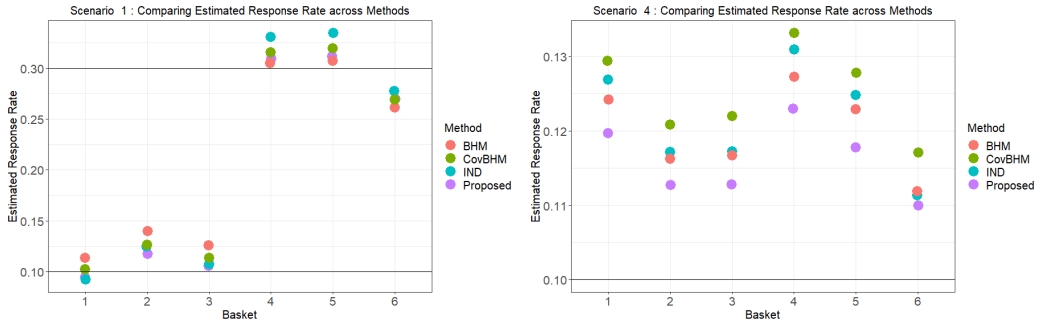
5. Simulation

We fit the binary response formulation of BPMx in (7) to simulated data to evaluate the model performance in diverse scenarios and compare with three other approaches: basket-wise independent analysis as in (1), a Bayesian hierarchical model (2) with the basket

Table 1: Description of the five scenarios for the simulation study. The $\{\alpha_b\}$ and $\{\xi_b\}$ are the basket-level means for y and x, respectively.

Basket	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5	
	α_b	ξ_b	α_b	ξ_b	α_b	ξ_b	α_b	ξ_b	α_b	ξ_b
1	0.1	10	0.1	10	0.05	10	0.1	10	0.1	10
2	0.1	10	0.1	10	0.05	10	0.1	10	0.1	10
3	0.1	10	0.1	10	0.2	15	0.1	10	0.1	10
4	0.3	18	0.3	10	0.2	15	0.1	18	0.1	10
5	0.3	18	0.3	10	0.35	20	0.1	18	0.1	10
6	0.3	18	0.3	10	0.35	20	0.1	18	0.1	10

Figure 1: The estimated basket-specific mean response probabilities for scenario 1. The four comparison methods are the Bayesian hierarchical model (BHM), the Bayesian hierarchical model with covariate (CovBHM), the independence model (IND), and the proposed BPMx. The horizontal lines are at the true mean response rates, 0.10 for baskets 1-3 and 0.30 for baskets 4-6.



parameters $\{\alpha_b\}$ arising from a common beta distribution, and a Bayesian hierarchical model with covariates (3) where the average covariate value for the basket, $\bar{x}_b$, linearly influences α_b on a logit scale. Priors on the model parameters were chosen so as to inject comparable prior information in these models. We assessed this by drawing samples from prior distributions and comparing them using empirical density curves.

We use the `Stan` language (Stan Development Team, 2021) to fit each of these models to obtain posterior samples using a Markov chain Monte Carlo sampling algorithm. We perform 50 replications for each scenario, with 5 chains of 1000 burn-in and 2000 inferential iterations. Data were generated for five scenarios, each with six baskets. The list of scenarios with true response probabilities and covariate means for every basket is in Table 1. Scenarios 1, 3 and 5 assume the same latent cluster structure for both response and covariate variables, with scenario 1 conforming to the two clusters anticipated by our model (7), and scenarios 3 and 5 having, respectively, three clusters and just one. Scenarios 2 and 4 violate the assumption of an informative covariate with same cluster structure as the response.

We estimate the basket-specific response rates, $\{\alpha_b\}$, using the posterior mean and estimate the precision of these estimates using the average root mean squared error.

Figure 1 shows the estimates of $\alpha_b, b = 1, 2, \ldots, 6$ in scenario 1 where the model

Figure 2: The average absolute bias in the estimation of the response means averaged across all baskets for simulation scenarios 1 and 4. The four comparison methods are the Bayesian hierarchical model (BHM), the Bayesian hierarchical model with covariate (CovBHM), the independence model (IND), and the proposed BPMx.

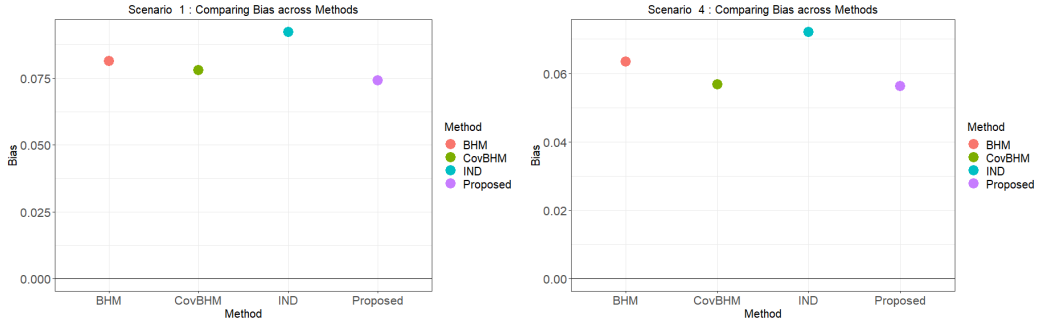
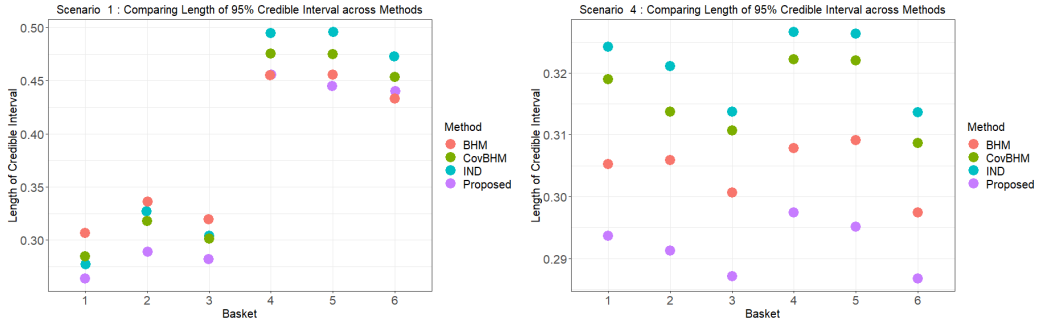


Figure 3: Average lengths of the response means averaged across all baskets for simulation scenario 1 and 4. The four comparison methods are the Bayesian hierarchical model (BHM), the Bayesian hierarchical model with covariate (CovBHM), the independence model (IND), and the proposed BPMx.



assumptions hold and in scenario 4 where the model assumptions regarding informative covariate and cluster structure are violated. BPMx was able to estimate the response probability well in all scenarios, including those where the assumptions regarding the underlying clusters are violated. BPMx also has the least overall absolute bias (see Figure 2) and root mean squared error (RMSE) (results not shown) across all scenarios. BPMx produced the shortest 95% credible intervals for response probabilities across most of the baskets for all the scenarios (see Figure 3).

We computed the efficacy and futility rates for all the baskets. In basket b, the treatment is concluded to be futile if $P(\alpha_b > 0.2 \mid data) < 0.05$, and effective if $P(\alpha_b > 0.1 \mid data) > 0.9$. The rates of efficacy and futility for every basket in scenarios 1 and 4 are in Figure 4. The BPMx was able to distinguish between responsive and non-responsive baskets well even when the underlying assumptions do not hold. Additionally, we also calculated the type 1 error rate as the percentage of times the treatment was concluded to be effective in a basket when it was in fact futile, and power as the percentage of times the treatment was correctly concluded to be effective in a basket. Using the same decision rules for all four models, Type 1 error rates were comparable, and the BPMx exhibited slightly higher power across all the scenarios (see Table 2 for results for scenarios 1 and 4).

There is an underlying cluster structure in the BPMx that is not present in the comparison models. The interpretation of the clusters as responsive and non-responsive may assist in making go/no-go decisions. Table 3 reports the MAP estimators of the partition among the 50 replications for each scenario, together with their relative frequency. In scenarios 1

Table 2: Type 1 error (T1E) and power for simulation scenarios 1 and 4. The four comparison methods are the Bayesian hierarchical model (BHM), the Bayesian hierarchical model with covariate (CovBHM), the independence model (IND), and the proposed BPMx.

Scenario	Basket b	True α_b	Proposed		IND		BHM		covBHM	
			T1E	Power	T1E	Power	T1E	Power	T1E	Power
1	1	0.10	0.06	-	0.06	-	0.06	-	0.06	-
	2	0.10	0.06	-	0.06	-	0.06	-	0.06	-
	3	0.10	0.04	-	0.04	-	0.06	-	0.04	-
	4	0.30	-	0.68	-	0.66	-	0.66	-	0.66
	5	0.30	-	0.68	-	0.66	-	0.66	-	0.66
	6	0.30	-	0.58	-	0.52	-	0.52	-	0.52
4	1	0.10	0.08	-	0.08	-	0.08	-	0.08	-
	2	0.10	0.08	-	0.08	-	0.08	-	0.08	-
	3	0.10	0.08	-	0.08	-	0.08	-	0.08	-
	4	0.10	0.12	-	0.12	-	0.12	-	0.12	-
	5	0.10	0.1	-	0.1	-	0.1	-	0.1	-
	6	0.10	0.02	-	0.02	-	0.02	-	0.02	-

Figure 4: Efficacy and futility rates for simulation scenarios 1 and 4. The four comparison methods are the Bayesian hierarchical model (BHM), the Bayesian hierarchical model with covariate (CovBHM), the independence model (IND), and the proposed BPMx.

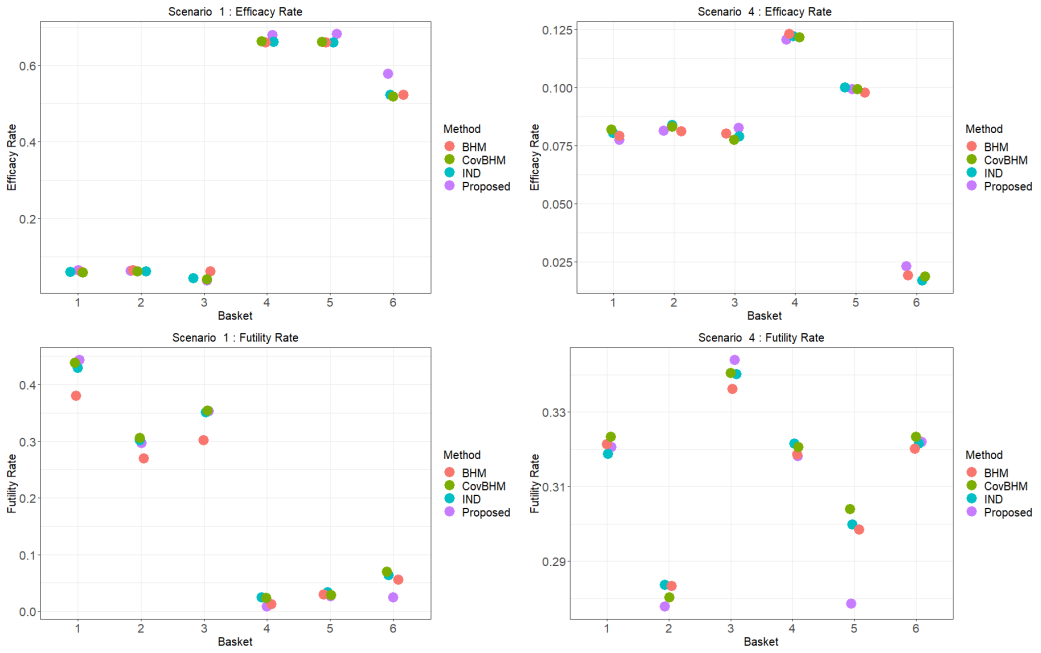


Table 3: MAP estimators of the partition among the 50 replications for each scenario, together with their relative frequency.

Scenario	True Partition	Estimated Partition	Rel. Frequency
1	$(1,1,1,2,2,2)$	$(1,1,1,2,2,2)$	1
2	$(1,1,1,2,2,2)$	$(1,1,1,1,1,1)$	1
		$(1,1,1,1,1,1)$	0.14
3	$(1,1,1,1,2,2)$	$(1,1,1,1,2,2)$	0.54
		$(1,1,1,2,2,2)$	0.02
		$(1,1,2,2,2,2)$	0.30
4	$(1,1,1,1,1,1)$	$(1,1,1,2,2,2)$	1
5	$(1,1,1,1,1,1)$	$(1,1,1,1,1,1)$	1

and 5 when the response and covariate have the same cluster structure, the correct partition was estimated with probability 1, meaning that BPMx identified the true partition in all replications. In scenario 3, when there are three underlying clusters in the data, the true partition was estimated with highest probability. However, BPMx failed to estimate the true partition in 46% of the replications. This could be due to the fact that the response probabilities in the three clusters are relatively close, which makes it more difficult to distinguish the three clusters. In scenarios 2 and 4 when the covariate is not truly informative of the treatment response, BPMx failed to identify the true partition. The estimated partition in this case was informed by the covariate.

We also estimate the cluster membership probability for each basket, $P_{bj}, b = 1, 2, .., B$, $j = 1, 2$. As stated earlier, in order to uniquely identify partitions, the first basket is always assigned to cluster 1, and the other cluster is called cluster 2. Hence a cluster is interpreted as baskets that have similar treatment response as opposed to strictly "responsive" and "non-responsive" baskets, keeping in mind that the first basket is always assigned to the first cluster. The cluster membership probabilities can be helpful in understanding the degree of similarity among the baskets.

Table 4 gives the estimated cluster membership probabilities. These probabilities indicate that BPMx was correctly able to identify the underlying number of latent clusters. As the cluster membership of basket 1 is fixed in the first cluster, the posterior probability estimate $\hat{P}_{11}$ is always estimated to be 1, making $\hat{P}_{12} = 1 - \hat{P}_{11} = 0$. The baskets with mean response similar to basket 1 were classified in the first cluster with high probability, whereas other baskets were classified to cluster 2 with high probability. In scenarios 1 and 5 where the response and covariate convey the same information, the baskets with similar response probabilities were identified correctly. In scenario 3 the covariate is informative of the response, but the number of clusters is more than 2. The baskets with intermediate response probabilities are assigned to both the clusters with probabilities close to 0.5. Thus, the estimated cluster membership probabilities are evident of the three clusters in the data. In scenarios 2 and 4 where the covariate is non-informative, the estimated cluster membership probabilities fail to identify the baskets with similar response probabilities correctly. This could be due to the possibility that these estimates are heavily influenced by the covariate information, as was the estimated partition for these scenarios.

6. Conclusion

BPMx can account for non-exchangeability through the assumption of latent clusters and the use of individual covariate information in partition models to estimate this latent

Table 4: Estimated cluster membership probabilities for each of the six baskets in each simulation scenario.

Scenario	Basket (b)	$\hat{P}_{b1}$	$\hat{P}_{b2}$
1	1	1	0
	2	0.98	0.02
	3	0.98	0.02
	4	0.04	0.96
	5	0.04	0.96
	6	0.04	0.96
2	1	1	0
	2	0.91	0.09
	3	0.91	0.09
	4	0.89	0.11
	5	0.89	0.11
	6	0.89	0.11
3	1	1	0
	2	0.94	0.06
	3	0.54	0.46
	4	0.54	0.46
	5	0.25	0.75
	6	0.25	0.75
4	1	1	0
	2	0.98	0.02
	3	0.98	0.02
	4	0.06	0.94
	5	0.06	0.94
	6	0.06	0.94
5	1	1	0
	2	0.9	0.1
	3	0.9	0.1
	4	0.9	0.1
	5	0.9	0.1
	6	0.9	0.1

cluster structure in the data and the treatment response. The simulation demonstrates that BPMx can estimate the mean treatment response with higher precision than the comparison models and these estimates are not degraded even when the underlying assumptions regarding the latent cluster structure are violated. However, in such a case, BPMx is unable to correctly estimate the partitions and cluster membership probabilities. When an informative covariate is used in the model, BPMx offers additional insights in terms of estimates of latent clusters and cluster membership probabilities for every basket.

Our belief that the covariate provides information about a subject's response to treatment underlies the assumption of the BPMx of the same latent cluster structure for both the response and covariate. The inability of BPMx to detect the true cluster structure under failure of this assumption may be due to the higher amount of information in the continuous covariate as compared to the binary response. More simulations, including for a formulation of the BPMx for a continuous response and continuous covariate, may help in understanding this further. Fixing the number of latent clusters to two, although restrictive, offers advantage in terms of interpretation. Increasing this number is a straight forward extension of the approach presented here, but it can pose more computational challenge. An informative prior specification on the partitions is another possible extension. Use of multiple covariates is also worth exploring further, with careful consideration towards the choice of similarity function.

Acknowledgements

The authors thank Dr. Kentaro Takeda, Dr. Shufang Liu and Dr. Alan Rong of Astellas Pharma, Northbrook, IL for their valuable guidance and support, and Lexi René, Florida State University for her helpful insights. They also acknowledge funding from NIH/NIMH R01MH121364.

References

Barry, D. and Hartigan, J. A. (1992). Product Partition Models for Change Point Problems. *The Annals of Statistics*, 20(1):260–279.

Berry, S. M., Broglio, K. R., Groshen, S., and Berry, D. A. (2013). Bayesian hierarchical modeling of patient subpopulations: Efficient designs of Phase II oncology clinical trials. *Clinical Trials*, 10(5):720–734. PMID: 23983156.

Chu, Y. and Yuan, Y. (2018a). A Bayesian basket trial design using a calibrated Bayesian hierarchical model. *Clinical Trials*, 15(2):149–158. PMID: 29499621.

Chu, Y. and Yuan, Y. (2018b). BLAST: Bayesian Latent Subgroup Design for Basket Trials Accounting for Patient Heterogeneity. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67:723–740.

Dahl, D. B. (2009). Modal clustering in a class of product partition models. *Bayesian Analysis*, 4(2):243 – 264.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457 – 472.

Hobbs, B. P. and Landin, R. (2018). Bayesian basket trial design with exchangeability monitoring. *Statistics in Medicine*, 37(25):3557–3572.

Hyman, D. M., Puzanov, I., Subbiah, V., Faris, J. E., Chau, I., Blay, J.-Y., Wolf, J., Raje, N. S., Diamond, E. L., Hollebecque, A., Gervais, R., Elez-Fernandez, M. E., Italiano, A., Hofheinz, R.-D., Hidalgo, M., Chan, E., Schuler, M., Lasserre, S. F., Makrutzki, M., Sirzen, F., Veronese, M. L., Tabernero, J., and Baselga, J. (2015). Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. *New England Journal of Medicine*, 373(8):726–736. PMID: 26287849.

Liu, R., Liu, Z., Ghadessi, M., and Vonk, R. (2017). Increasing the efficiency of oncology basket trials using a Bayesian approach. *Contemporary Clinical Trials*, 63:67– 72.

Müller, P., Quintana, F., and Rosner, G. (2011). A Product Partition Model With Regression on Covariates. *Journal of Computational and Graphical Statistics : A Joint Publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*, 20:260–278.

Neuenschwander, B., Wandel, S., Roychoudhury, S., and Bailey, S. (2016). Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharmaceutical Statistics*, 15(2):123–134.

Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials*, 10:1 – 10.

Stan Development Team (2021). *Stan Modeling Language Users Guide and Reference Manual*.

Thall, P., Kyle Wathen, J., Nebiyou Bekele, B., Champlin, R., Baker, L., and S Benjamin, R. (2003). Hierarchical Bayesian approaches to phase II trials in disease with multiple subtypes. *Statistics in Medicine*, 22:763–80.

Ventz, S., Barry, W. T., Parmigiani, G., and Trippa, L. (2017). Bayesian response-adaptive designs for basket trials. *Biometrics*, 73(3):905–915.

Zhou, T. and Ji, Y. (2020). RoBoT: a robust Bayesian hypothesis testing method for basket trials. *Biostatistics*.