ASID: A Bayesian Adaptive Subgroup-Identification Enrichment Design

Florica Constantine[*], Yili Pritchett[†], Yanxun Xu[*‡]

Abstract

Targeted therapies based on patients' baseline characteristics such as biomarkers have been growing interests for many diseases. Depending on the expression of specific biomarkers or their combinations, different patient subgroups could respond differently to the same treatment. An ideal design, especially at the proof of concept stage, should search for such subgroups and make dynamic adaptation as the trial goes on. When no prior knowledge is available on whether the treatment works on the all-comer population or only works on the subgroup defined by one biomarker or several biomarkers, it is necessary to estimate the subgroup effects adaptively based on the response outcomes and biomarker profiles from all the treated subjects at the interim analysis. To address this problem, we propose an adaptive subgroup-identification enrichment design, ASID, which can simultaneously search for predictive biomarkers, estimate the subgroups with differential treatment effects, and modify the study entry criteria at the interim analysis. We compare the ASID with an alternative adaptive enrichment design based on linear regression in a motivating Alzheimer's disease clinical trial, and demonstrate via simulation the superior performance of the ASID.

Key Words: Alzheimer's disease, Biomarker, Clinical trial, Targeted therapies

1. Introduction

1.1 Background

For many diseases, it has become well known that there exists a heterogeneity of treatment effects across patient subpopulations that depend on the expression of specific biomarkers or their combinations when given the same treatment. Thus, it is essential in clinical trial designs or clinical data analysis to take into account the potentially different treatment effects among patient subpopulations when making a decision. For instance, breast cancer patients with an enriched HER2 pathway were found to respond well to the medication trastuzumab (Hudis, 2007) through pairing genetic traits with targeted treatment options, while other subtypes of breast cancers do not. Another example is that treatments with EGFR antibodies are not recommended for KRAS mutated colorectal cancer patients since they are usually resistant to anti-EGFR treatment (Misale et al., 2012). Therefore, it is very important to identify the biomarkers that have interaction effects with the treatment and are predictive of the subgroups that are more likely to respond to the treatment.

For situations where the predictive biomarkers are not clear based on the treatment's mechanism or disease clinical presentation, adaptive enrichment designs have been developed to adaptively modify the eligibility rules at interim analyses in an effort to determine the appropriate biomarkers. Patients exhibiting the desired treatment effects are referred to as the "enriched population". Karuri and Simon (2012) compared the treatment versus placebo by evaluating the treatment effectiveness of biomarker positive- and negative-subgroups at an interim analysis, allowing the termination of enrollment of the biomarker negative subgroup. Simon and Simon (2013) developed a class of adaptive enrichment designs to adaptively update the eligibility criteria, while preserving the type I error. See,

[*]Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD.

[†]MedImmune, LLC, Gaithersburg, MD.

[‡]Correspondence should be addressed to Yanxun Xu (yanxun.xu@jhu.edu)

for example, Wang and Hung (2013) for more extensive reviews of adaptive enrichment design.

Most of these enrichment designs use a set of biomarkers to pre-define subgroups and then test if there are differential therapeutic effects on these pre-defined subgroups. However, pre-defining subgroups can be problematic if the pre-defined subgroups are not predictive of patients' responses or treatment selections. Therefore, an enrichment design that allows the discovery and estimation of the subgroups during the clinical trial is desired. Sivaganesan et al. (2011) identified subgroups through a model selection procedure with each biomarker defining classes of models. Foster et al. (2011) developed a random forest-based algorithm to find the subgroups by searching biomarker regions where the treatment effect is larger than the average effect on the whole population. Loh et al. (2015) proposed a regression tree approach, GUIDE, to first decide which biomarkers to split on through the use of χ^2 tests, and then to identify subgroups with differential treatment effects. Shen and He (2015) used a structured logistic-normal mixture model to test for the existence of subgroups by a confirmatory statistical test. All of these methods target the subgroup identification using the retrospective clinical trial data and have not been directly applied to the clinical trial designs.

There have been relatively few methodologies on adaptively estimating the subgroups based on patients' differential responses to treatments during interim analyses. Xu et al. (2014) developed SUBA, a Bayesian subgroup-based adaptive design, to allocate the patients to their superior treatments using a random partition model that splits the biomarker space by the observed biomarker's median value, which generally is not the optimal cutoff for a predictive biomarker. Guo et al. (2016) extended SUBA (*i.e.*, SCUBA) by allowing the biomarker space to be split using hyperplanes that construct linear boundaries, providing a more flexible partition model. However, both SUBA and SCUBA focus on the subgroup identification for patient allocations and do not modify the study entry criteria during the trial.

1.2 A Biomarker-Driven and Subgroup-Based Enrichment Design

In this paper, we propose a class of adaptive subgroup-identification enrichment designs (ASID), utilizing patients' biomarker profiles and outcomes as they become available. ASID searches for subgroups among a set of biomarkers rather than predefining subgroups and allows the entry criteria to be modified to enroll more subjects who are more likely to respond to the treatment, improving the chance of detecting a clinically relevant treatment effect. ASID incorporates a flexible Bayesian model that can handle covariates of varying forms (continuous, binary, categorical, or ordinal) and different types of outcomes (binary, categorical, continuous, survival).

The key innovation of ASID is that the subgroups with differential treatment effects are continuously identified, redefined and ranked based on patients' responses data at each interim analysis using a pre-specified algorithm where Bayesian models are incorporated. The continuous search for a group with a differential treatment effect will lead to one time modification of study enrollment criteria. Therefore, more patients with the characteristics as in the identified subgroup can be enrolled to the study as the trial goes on. This design combines the task of subgroup identification and that of ensuring adequate sample size to learn the investigational compound's efficacy and safety in the identified sub-population. This is superior to the conventional approach where either two clinical trials are needed to accomplish these tasks, or the sample size is not enough for the post-hoc identified subgroup to inform the next step of development for the compound.

1.3 Motivating Trial

We consider a placebo-controlled, double-blind proof-of-concept (POC) study for an investigational compound on patients with Alzheimer's disease. The primary efficacy endpoint is the change from baseline to final observation on the 13-item Alzheimer's disease (AD) Assessment Scale – Cognitive Subscale (ADAS-cog) score. Previous research has suggested a number of biomarkers that might predict treatment effect on AD, and these biomarkers are apolipoprotein E (APOE)-ϵ4 genotype and allele status, plasma amyloid precursor protein β (Aβ), and cerebrospinal fluid (CSF) β-site amyloid precursor protein (APP)-cleaving enzyme 1 (BACE1). It is plausible that the investigational compound only has a clinically meaningful effect on a sub-population that is qualified by one of the biomarkers or a combination of several biomarkers listed above. The objective of this clinical trial is to combine the task of subgroup identification and population enrichment to gain efficiency with the purpose of developing this investigational compound more rapidly.

In this study, patients with AD meeting entry criteria will be enrolled and equally assigned to placebo or the investigational compound. Baseline biomarkers data will be collected. At the pre-specified interim analysis, accumulating ADAS-cog total scores are utilized to search for a potential subgroup with differential treatment effect following a pre-specified algorithm. When such a subgroup is ascertained, study entry criteria will be modified so that only the patients with the characteristics in the identified sub-population will be enrolled to the study. This design allows population enrichment in the middle course so that more information can be obtained for the sub-population to inform the next step of clinical development.

This paper proceeds as follows. We first introduce the Bayesian probability model in Section 2. The proposed ASID design is summarized in Section 3. Section 4 presents extensive simulation studies and examines the operating characteristic of the ASID. Finally, we conclude with a discussion in Section 5.

2. Probability Model

2.1 Sampling Model

Assume we have a maximum sample size of N patients that are indexed by $i = 1, \ldots, N$, and suppose there are T candidate treatments indexed by $t = 1, \ldots, T$. In the motivating AD trial, $T = 2$. Let $z_i = t$ denote that patient i is assigned to treatment t. There are K biomarkers that are identified from the investigational compound's mechanism or disease clinical presentation. We assume that a biomarker k can be binary, ordinal, categorical, or continuous, where $k = 1, \ldots, K$. Denote $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iK})'$ and y_i to be the biomarker profile and the response outcome of the i^{th} patient, respectively.

Denote Ω to be the biomarker space. We say that a partition is a family of subsets $\Pi = \{S_1, \cdots, S_m, \cdots, S_M\}$, where the S_m's are mutually disjoint and their union is Ω. Here the number of subsets M is random. The partition on the biomarker space induces a partition of the patients. If $\boldsymbol{x}_i \in S_m$, we say patient i with biomarker profile $\boldsymbol{x}_i$ belongs to subgroup m. We will construct a prior probability measure for Π in the next section. Below we consider the sampling model of y_i of different types conditional on $\boldsymbol{x}_i$ and Π. Let Θ denote the parameters in the sampling model. Denote $\boldsymbol{Y}_n = (y_1, \ldots, y_n)$, $\boldsymbol{X}_n = \{\boldsymbol{x}_i\}_{i=1}^n$, and $\boldsymbol{Z}_n = (z_1, \ldots, z_n)$.

Binary outcomes. Let $y_i \in \{0, 1\}$ and $\theta_{t,m}$ be the response rate of patients in subgroup m under treatment t. In this case, $\Theta = \{\theta_{t,m}\}$. We assume $p(y_i = 1 \mid z_i = t, \Pi, \boldsymbol{x}_i \in S_m) = \theta_{t,m}$. The likelihood function is simply the product of n Bernoulli probability mass

functions. We assign the prior $\theta_{t,m} \mid \Pi \overset{iid}{\sim} Beta(a,b)$, where $Beta(a,b)$ denotes a beta distribution with mean $a/(a+b)$.

Categorical outcomes. Let $y_i \in \{1,2,\ldots,C\}$, where C is the number of outcome categories. Denote $\theta_{c,t,m}$ to be the response rate of patients in subgroup m under treatment t for outcome category c. We assume $p(y_i = c \mid z_i = t, \Pi, \boldsymbol{x}_i \in S_m) = \theta_{c,t,m}$, where $\sum_{c=1}^{C} \theta_{c,t,m} = 1$. We assign the conjugate prior

$$(\theta_{1,t,m},\ldots,\theta_{C,t,m}) \mid \Pi \overset{iid}{\sim} \text{Dirichlet}(a_1,\ldots,a_C).$$

Continuous outcomes. Let $y_i \in \mathbb{R}$ and $\theta_{t,m}$ be the mean response of patients in subgroup m under treatment t. We assume $p(y_i \mid z_i = t, \Pi, \boldsymbol{x}_i \in S_m) = N(\theta_{t,m}, \sigma^2)$. The likelihood can be written as follows:

$$p(\boldsymbol{Y}_n \mid \boldsymbol{X}_n, \boldsymbol{Z}_n, \Theta, \Pi) = \prod_{t=1}^{T}\prod_{m=1}^{M}\prod_{\{i:z_i=t,\boldsymbol{x}_i\in S_m\}} (2\pi\,\sigma^2)^{-1/2}\exp\{-\frac{1}{2\sigma^2}(y_i-\theta_{t,m})^2\}. \tag{2.1}$$

We assign the conjugate prior $p(\theta_{t,m},\sigma^2) = p(\theta_{t,m}|\sigma^2)p(\sigma^2)$ with

$$p(\theta_{t,m}|\sigma^2) = N(\theta_0, \frac{\sigma^2}{\kappa_0}) \text{ and } p(\sigma^2) = IG(\frac{\nu_0}{2}, \frac{SS_0^2}{2}),$$

where $SS_0^2 = \nu_0\sigma_0^2$.

Survival outcomes. We assume that $\log(y_i) = \boldsymbol{x}_i'\boldsymbol{\beta}_{t,m} + \epsilon_i$, where $\boldsymbol{\beta}_{t,m}$ denotes the regression coefficient for patients in subgroup m under treatment t and $\epsilon_i \sim N(0,\sigma^2)$. We assign the priors $\boldsymbol{\beta}_{t,m} \sim N(\mu_0,\Sigma_0)$ and $\sigma^2 = IG(a_0,b_0)$.

The joint model can be written as follows,

$$p(\boldsymbol{Y}_n, \Theta, \Pi \mid \boldsymbol{X}_n, \boldsymbol{Z}_n) \propto p(\boldsymbol{Y}_n \mid \boldsymbol{X}_n, \boldsymbol{Z}_n, \Theta, \Pi)\, p(\Theta \mid \Pi)\, p(\Pi \mid \boldsymbol{c})p(\boldsymbol{c}), \tag{2.2}$$

where $\boldsymbol{c}$ denotes the parameters in the model that describes the random partition Π. We have introduced the sampling model and the priors for Θ. In the next section, we will discuss the prior of Π.

2.2 Prior of Partition Π

We propose a tree-type random partition on the biomarker space Ω to define random biomarker subgroups. We build partitions via a tree of recursive splits: each node of the tree represents a subset of Ω. The final leaves of the tree are the partitioning sets S_m. At each node the tree is either pruned or the corresponding subset is further split into two siblings. In the second case, the two siblings are defined by a plane orthogonal to a randomly selected axis of Ω, say the axis of the k-th biomarker. In other words, through a sequence of splits, each of which selects a biomarker k first and then splits the space of x_k in half, we generate a partition set of Ω as the collection of the resulting subsets. For the motivating AD trial, we limit the partition to at most four biomarker subgroups, and hence no more than two rounds of random splits in the random partition. This constraint is imposed to limit the number of subgroups with very few patients.

Figure 1 illustrates a simple case with two rounds of splits and two continuous biomarkers on $[-1,1]^2$. In each round, for each of the current subsets, we split along a biomarker k with probability ν_k or choose not to split with probability ν_0, $\sum_{k=0}^{K}\nu_k = 1$. If an ancestor subset S is split into two subsets by the k^{th} biomarker, then the resulting subsets are $\{i : x_{ik} < Th_k(S)\}$ and $\{i : x_{ik} \geq Th_k(S)\}$, where $Th_k(S)$ is the threshold by which

the subset is being split. For example, in Figure 1, biomarker 1 is chosen in the first round and the patients are split into $U_1 = \{i : x_{i1} < 0.5\}$ and $L_1 = \{i : x_{i1} \geq 0.5\}$. Here U and L denote that the measurements are smaller and larger than the threshold, respectively. In round 2, we split subgroup U_1 into two new biomarker subgroups UU_{11} and UL_{11} by choosing biomarker 1 with threshold 0 and split biomarker subgroup L_1 into two new biomarker subgroups LU_{12} and LL_{12} by choosing biomarker 2 with threshold -0.2. Note that the ordering of letters U and L are matched with the ordering of the biomarker index. Therefore, at the end, the partition $\Pi = \{UU_{11}, UL_{11}, LU_{12}, LL_{12}\}$, which corresponds to four biomarker subgroups.

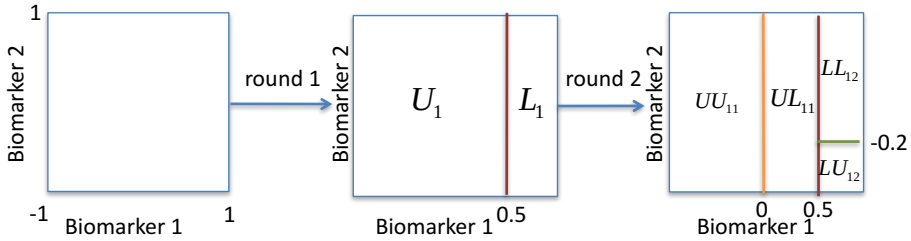

Figure 1: An illustration of Π. The example shows that with two rounds of split, the initial space of two biomarkers is partitioned into four subsets $\{UU_{11}, UL_{11}, LU_{12}, LL_{12}\}$.

In the Appendix, we describe the detailed split rules using two rounds of splits by taking into account various types of covariates, including continuous, binary, categorical, and ordinal variables.

3. Trial Design

In the following discussion, we assume that the outcome is continuous, $T = 2$, and one interim analysis is needed, as in the motivating AD trial. However, the ASID can be easily extended to multiple treatments and multiple interim analyses. At the interim analysis, we would like to discover the regions in the biomarker space Ω in which the patients will benefit from the treatment. Since a random distribution is proposed on the partition, summarizing a distribution over random partitions and discovering the subgroups become challenging. To address this problem, we report the subgroups with differential treatment effects as follows.

Assume n patients have been treated and their responses have been observed before the interim analysis. Denote $\mathcal{D}_n = \{\boldsymbol{Y}_n, \boldsymbol{X}_n, \boldsymbol{Z}_n\}$. In the biomarker space Ω, we select D_k grid points on biomarker k, and then take the Cartesian product of the grids across all K biomarkers. For example, if biomarker k is continuous on [-1, 1], we can choose $D_k = 20$ equally spaced points on [-1, 1]; if biomarker k is binary, then $D_k = 2$. Each grid point d represents a possible biomarker profile. In the MCMC samples, the b^{th} iteration after burn-in generates a posterior sample $\{\Theta^{(b)}, \Pi^{(b)}, \boldsymbol{c}^{(b)}\}$, which defines a partition set $\Pi^{(b)} = \{S_1^{(b)}, \ldots, S_{M^{(b)}}^{(b)}\}$ and their corresponding response parameters. For each grid point d with biomarker profile $\boldsymbol{x}_d$, we can find the subgroup $S_m^{(b)}$ it belongs to, and the response parameter $\theta_{t,d}^{(b)} = \theta_{t,m}^{(b)}$ if $z_i = t$ and $\boldsymbol{x}_d \in S_m^{(b)}$.

In the motivating AD trial, we assume that $t = 0$ represents placebo and $t = 1$ represents the investigational compound. Denote $q_d^{(b)} = \theta_{1,d}^{(b)} - \theta_{0,d}^{(b)}$, then the posterior probability of the treatment effect of grid d with biomarker $\boldsymbol{x}_d$ larger than the low reference value (LRV) can be computed as $\delta_d = \frac{1}{B}\sum_{b=1}^{B} I\{q_d^{(b)} \geq \text{LRV}\}$. Here LRV represents a

clinically meaningful minimum increment and B is the number of MCMC iterations after burn-in.

The ASID will be conducted as follows.

- Start the trial. The first n patients are equally randomized to the placebo and the investigational compound.
- At the interim analysis, if there exists a subgroup Δ that is a convex hull (Graham, 1972), the smallest convex set that contains $\{d : \delta_d \geq \xi\}$ for some threshold ξ, we restrict the entry into the clinical trial to only patients with $\boldsymbol{x} \in \Delta$. If $\Delta = \emptyset$, we stop the trial.
- If $\Delta \neq \emptyset$, we continue to recruit additional $N - n$ patients equally randomized to the placebo and the investigational compound with the biomarkers $\boldsymbol{x} \in \Delta$.

4. Simulation Studies

4.1 Simulation Setup

We conducted simulation studies based on the motivating AD trial to evaluate the performance of the proposed ASID design. In each trial, the maximum sample size was $N = 140$ patients and LRV $= 2.37$. We assumed that $K = 4$ baseline biomarkers were available for each patient and $p(\nu_k) = 1/5$, $k = 0, 1, \ldots, 4$, indicating a uniform prior on the biomarker selection. The priors on the parameters in $\boldsymbol{c}$ were introduced in the Appendix. As a run-in phase, $n = 80$ patients were equally randomized to the placebo and the investigational compound. In scenarios 1-3, we generated x_{ik} from $\mathrm{Uniform}(-1, 1)$, $i = 1, \ldots, n$ and $k = 1, \ldots, 4$. In scenario 4, we consider three continuous biomarkers generated from $\mathrm{Uniform}(-1, 1)$ and one binary biomarker generated from $\mathrm{Bernoulli}(0.5)$. For the proposed subgroup report, we fixed the parameter $\xi = 0.9$ based on extensive sensitivity analysis (not shown here).

For comparisons, we implemented an alternative design for each simulated trial, a linear regression (LR) design. Under the LR design, the outcomes were modeled as a Bayesian linear regression considering all main effects and the interaction effects between the treatment and the biomarkers: $y_i \mid z_i, \boldsymbol{x}_i = \beta_0 + \beta_1 z_i + \boldsymbol{\alpha}\boldsymbol{x}_i + \boldsymbol{\gamma} z_i \boldsymbol{x}_i + \epsilon_i$, where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. We assumed non-informative conjugate priors, $1/\sigma^2 \sim \mathrm{Gamma}(0.1, 0.1)$ and $(\beta_0, \beta_1, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \sim MN(0, 20I)$, where $\mathrm{Gamma}(a, b)$ denotes a gamma distribution with mean a/b and $MN(\boldsymbol{\mu}, \Sigma)$ denotes a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance Σ. The posterior samples were obtained by a Gibbs sampling procedure. We computed $q_d^{(b)} = E(y_i \mid z_i = 1, \boldsymbol{x}_d) - E(y_i \mid z_i = 0, \boldsymbol{x}_d)$.

We considered four scenarios and simulated 100 trials for each scenario. In scenario 1, we assumed $y_i = 0.75 + 0.25I(z_i = 1) + 3.5I(x_{i1} > -0.4)I(z_i = 1) + \epsilon_i$. In scenario 2, the outcomes y_i's are generated from $y_i = 0.75 + 0.25I(z_i = 1) + 3.5I(x_{i1} < 0.4, x_{i2} > -0.4)I(z_i = 1) + \epsilon_i$. In scenario 3, $y_i = 0.75 + 0.25I(z_i = 1) + 3.5I(x_{i1} = 1, x_{i2} > -0.4)I(z_i = 1) + \epsilon_i$. In scenario 4, we assumed $y_i = 0.75 + 0.25I(z_i = 1) + 1.5I(x_{i1} = 1, x_{i2} > -0.4)I(z_i = 1) + \epsilon_i$. Here $\epsilon_i \sim N(0, 1)$. Define the true effective subgroup as $S^o = \{i : [E(y_i \mid z_i = 1, \boldsymbol{x}_i) - E(y_i \mid z_i = 0, \boldsymbol{x}_i)] > \mathrm{LRV}\}$. The left column of Figure 2 show the simulated true effective subgroups in blue color for scenario 1-3. In scenario 4, the true effective subgroup is $\emptyset$.

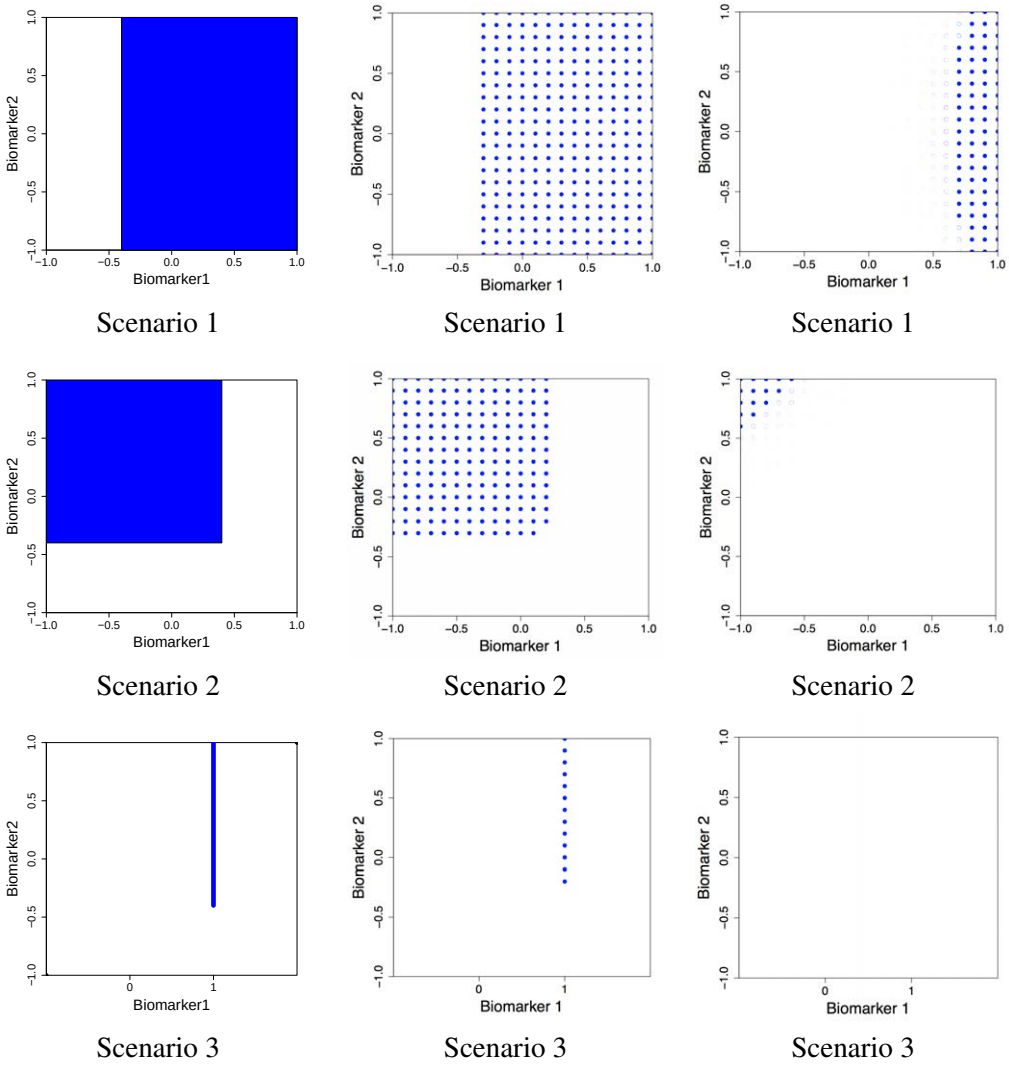

Figure 2: Left column: true effective subgroups in blue for scenarios 1-3 in the simulation. Middle column: the posterior estimated subgroups represented by grid points in blue under the ASID. Right column: the posterior estimated subgroups represented by grid points in blue under the LR design.

4.2 Simulation Results

We first report the subgroup finding at the interim analysis. Denote δ_d^h to be the posterior probability of the treatment effect of grid point d with biomarker profile $\boldsymbol{x}_d$ larger than LRV in trial h, $h = 1, \ldots, 100$. We computed $\hat{\Delta} = \{d : \frac{1}{100}\sum_{h=1}^{100} \delta_d^h \geq \xi\}$. The middle column of Figure 2 shows the posterior estimate subgroup $\hat{\Delta}$ under the ASID, represented by the grid points in blue for scenarios 1-3. As shown in Figure 2, the effective subgroups identified by the ASID have a large overlap with the simulation truth. Scenario 4 is a NULL case, there is no effective subgroup. The ASID identified $\hat{\Delta} = \emptyset$, which matches the simulation truth. The right column of Figure 2 plots the posterior estimated subgroup by the LR design, which performs much worse than the ASID. For scenario 4, the LR design also recovers the simulated truth.

In addition, we report the sensitivity (true positive rate) and specificity (true negative rate) of the subgroup finding as the operating characteristics of the ASID and the LR design:

1) True positive rate = $\sum_{\{d:\boldsymbol{x}_d \in S^o\}} \sum_{h=1}^{100} I(\boldsymbol{x}_d \in \hat{\Delta})/(|S^o| \times 100)$, where $|S^o|$ is the number of grid points in the simulated true effective subgroup; 2) True negative rate = $\sum_{\{d:\boldsymbol{x}_d \notin S^o\}} \sum_{h=1}^{100} I(\boldsymbol{x}_d \notin \hat{\Delta})/(|\Omega \setminus S^o| \times 100)$. As shown in Table 1(a), both the ASID and the LR design achieve high specificity, but the ASID achieves much higher sensitivity compared to the LR design.

	Sensitivity		Specificity	
Scenario	ASID	LR	ASID	LR
1	0.91	0.54	1.00	1.00
2	0.90	0.38	0.99	0.98
3	0.87	0.00	1.00	1.00

(a)

	OMT	
Scenario	ASID	WO
1	3.75	2.74
2	3.74	1.93
3	3.71	1.51

(b)

Table 1: (a): The operating characteristics of the ASID and the LR designs. (b): The overall mean treatment effect of the ASID and the WO design.

Lastly, we study the ASID efficacy by comparing the proposed enrichment design with the design without the enrichment procedure. That means, we do not modify the study entry criteria during the interim analysis. We call this design the WO design. Denote y_i^h and z_i^h to be the response and treatment assignment for patient i in the h^{th} simulated trial, $h = 1, \ldots, 100$, we define the overall mean treatment effect (OMT) as $\mathrm{OMT} = \frac{1}{100} \sum_{h=1}^{100} \big(\frac{\sum_{i=n+1}^{N} y_i^h I(z_i^h=1)}{\sum_{i=n+1}^{N} I(z_i^h=1)} - \frac{\sum_{i=n+1}^{N} y_i^h I(z_i^h=0)}{\sum_{i=n+1}^{N} I(z_i^h=0)} \big)$, which is the mean response differences between the investigational compound and the placebo after the interim analysis. As shown in Table 1(b), the ASID yields higher OMT compared to the WO design.

5. Conclusion

We demonstrated the importance of identifying subgroups and modifying the study entry criteria in an adaptive enrichment design when there exist subgroups with differential treatment effects. The key contributions of the ASID include the construction of the random partition model that allows a flexible algorithm to explore subgroups and the modification of study entry criteria at the interim analysis. As shown in the simulation studies, the ASID can successfully recover the simulated effective subgroups. Compared to the alternative LR design, the ASID achieves much better sensitivity.

The proposed ASID focuses on the study entry criteria at the interim analysis by identifying subgroups of patients with differential treatment effects. One could easily add to the ASID an adaptive patient allocation algorithm to assign patients to their superior treatments and a final recommendation of a suitable patient population for a follow-up trial. As one future research direction, when a large number of biomarkers are available, we can extend the ASID to incorporate variable selection.

Acknowledgement

This work is funded by the MedImmune L.L.C. and Johns Hopkins University collaboration project.

References

Foster, J. C., Taylor, J. M., and Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Statistics in Medicine*, 30(24):2867–2880.

Graham, R. L. (1972). An efficient algorith for determining the convex hull of a finite planar set. *Information Processing Letters*, 1(4):132–133.

Guo, W., Ji, Y., and Catenacci, D. V. (2016). A subgroup cluster-based bayesian adaptive design for precision medicine. *Biometrics*.

Hudis, C. A. (2007). Trastuzumab—mechanism of action and use in clinical practice. *New England Journal of Medicine*, 357(1):39–51.

Karuri, S. W. and Simon, R. (2012). A two-stage Bayesian design for co-development of new drugs and companion diagnostics. *Statistics in Medicine*, 31(10):901–914.

Loh, W.-Y., He, X., and Man, M. (2015). A regression tree approach to identifying subgroups with differential treatment effects. *Statistics in Medicine*, 34(11):1818–1833.

Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G., et al. (2012). Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature*, 486(7404):532–536.

Shen, J. and He, X. (2015). Inference for subgroup analysis with a structured logistic-normal mixture model. *Journal of the American Statistical Association*, 110(509):303–312.

Simon, N. and Simon, R. (2013). Adaptive enrichment designs for clinical trials. *Biostatistics*, page kxt010.

Sivaganesan, S., Laud, P. W., and Müller, P. (2011). A Bayesian subgroup analysis with a zero-enriched Polya Urn scheme. *Statistics in Medicine*, 30(4):312–323.

Wang, S.-J. and Hung, H. J. (2013). Adaptive enrichment with subpopulation selection at interim: methodologies, applications and design considerations. *Contemporary Clinical Trials*, 36(2):673–681.

Xu, Y., Trippa, L., Müller, P., and Ji, Y. (2014). Subgroup-based adaptive (SUBA) designs for multi-arm biomarker trials. *Statistics in Biosciences*, pages 1–22.

Appendix

Below we describe the split rules for determining a partition of the biomarker space using the example of two rounds of splits and various types of covariates, including continuous, binary, categorical, and ordinal variables.

1. In the first split, we select biomarker k with probability ν_k, $k = 1, \ldots, K$, and choose to split or not to split with probability ν_0. We assume $\nu_k = \frac{1}{K+1}$, indicating a uniform prior. Then we choose threshold c_k to split the biomarker space into two subgroups U_k and L_k. The prior of $p(c_k)$ differs depending on what type of variable the biomarker is.

   (a) If biomarker k is binary, the split will be deterministic and we denote $U_k = \{i : x_{ik} = 0\}$ and $L_k = \{i : x_{ik} = 1\}$. Therefore $p(c_k) = 1$.

   (b) If biomarker k is continuous, denote $U_k = \{i : x_{ik} \leq c_k\}$ and $L_k = \{i : x_{ik} > c_k\}$. We assume $p(c_k) = \text{Uniform}(\min\{x_{ik}\}_{i=1}^N, \max\{x_{ik}\}_{i=1}^N)$.

   (c) If biomarker k is ordinal, let V_k denote the number of labels that biomarker k has. Let c_k denote the endpoint of the left partition, e.g., if $V_k = 5$ and $c_k = 3$, the left partition is $\{1, 2, 3\}$ and the right partition is $\{4, 5\}$. In this way we denote $U_k = \{i : x_{ik} \leq c_k\}$ and $L_k = \{i : x_{ik} > c_k\}$. Moreover, if $c_k = V_k$, it is equivalent to not splitting, which has been considered with probability ν_0. Therefore, we assume that $p(c_k) = \frac{1}{V_k - 1}$.

   (d) If biomarker k is categorical, let V_k denote the number of categories corresponding to biomarker k. Let c_k denote the elements in one subset U_k. The remaining elements are stored in the other subset L_k. The c_k are elements of the powerset of $\{1, 2, \cdots, V_k\}$ without the empty-set or the full set. There are hence $2^{V_k} - 2$ options for c_k. Note that the choice of c_k is symmetric: we may flip c_k and its complement while also exchanging k_1 and k_2, the biomarkers chosen to split on in the second round of splits, without any change. Thus, $p(c_k) = \frac{2}{2^{V_k} - 2}$.

2. In the second split, in the subset U_k, we choose biomarker k_1 with probability $p(k_1) = \frac{1}{K+1}$ since we can split along any biomarker k_1 or choose not to split. Then we choose threshold c_{k_1} to split the subgroup U_k into two subgroups UU_{kk_1} and UL_{kk_1}. In the subset L_k, we choose biomarker k_2 with probability $p(k_2) = \frac{1}{K+1}$ since we can split along any biomarker k_2 or choose not to split. Then we choose threshold c_{k_2} to split the subgroup L_k into two subgroups LU_{kk_2} and LL_{kk_2}. Here, $p(c_{k_1})$ and $p(c_{k_2})$ differ depending on what type of variable the biomarker is and the values obtained for c_{k_1} and c_{k_2}.

   (a) If biomarker k_1 is continuous, denote $UU_{kk_1} = \{i : \boldsymbol{x}_i \in U_k \text{ and } x_{ik_1} \leq c_{k_1}\}$ and $UL_{kk_1} = \{i : \boldsymbol{x}_i \in U_k \text{ and } x_{ik_1} > c_{k_1}\}$. We have $p(c_{k_1}) = \frac{1}{2}$ if $k_1 \neq k$; $p(c_{k_1}) = \frac{1}{c_k + 1}$ if $k_1 = k$.
   If biomarker k_2 is continuous, denote $LU_{kk_2} = \{i : \boldsymbol{x}_i \in L_k \text{ and } x_{ik_1} \leq c_{k_1}\}$ and $LL_{kk_2} = \{i : \boldsymbol{x}_i \in L_k \text{ and } x_{ik_1} > c_{k_1}\}$. We have $p(c_{k_2}) = \frac{1}{2}$ if $k_2 \neq k$; $p(c_{k_2}) = \frac{1}{1 - c_k}$ if $k_2 = k$.

   (b) If c_{k_1} is binary, denote $UU_{kk_1} = \{i : \boldsymbol{x}_i \in U_k, \text{ and } x_{ik_1} = 0\}$ and $UL_{kk_1} = \{i : \boldsymbol{x}_i \in U_k, \text{ and } x_{ik_1} = 1\}$ with $p(c_{k_1}) = 1$.

If c_{k_2} is binary, denote $LU_{kk_2} = \{i : \boldsymbol{x}_i \in L_k, \text{ and } x_{ik_2} = 0\}$ and $LL_{kk_2} = \{i : \boldsymbol{x}_i \in L_k, \text{ and } x_{ik_2} = 1\}$ with $p(c_{k_2}) = 1$.

(c) If c_{k_1} is ordinal, we let c_{k_1} denote the left endpoint of the second split *within* the left partition, and if c_{k_2} is ordinal, we let c_{k_2} analogously denote the left endpoint *within* the right partition. We have $p(c_{k_1}) = \frac{1}{V_{k_1}-1}$ if $k_1 \neq k$; $p(c_{k_1}) = \frac{1}{c_k}$ if $k_1 = k$. Also, $p(c_{k_2}) = \frac{1}{V_{k_2}-1}$ if $k_2 \neq k$; $p(c_{k_2}) = \frac{1}{V_k - c_k}$ if $k_2 = k$.

(d) If c_{k_1} is categorical, we let c_{k_1} denote one subset of the split *within* the subset U_k.
If c_{k_2} is categorical, we let c_{k_2} denote one subset of the split *within* the subset L_k. We have $p(c_{k_1}) = \frac{2}{2^{V_{k_1}}-2}$ if $k_1 \neq k$; $p(c_{k_1}) = \frac{2}{2^{|c_k|}-2}$ if $k_1 = k$. Also, $p(c_{k_2}) = \frac{2}{2^{V_{k_2}}-2}$ if $k_2 \neq k$; $p(c_{k_2}) = \frac{2}{2^{V_k - |c_k|}-2}$ if $k_2 = k$.