

Adaptive Signature Design- review of the biomarker guided adaptive phase –III controlled design

Arinjita Bhattacharyya^a, Shesh N. Rai^{a,b,c,*}

^a Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, KY, USA

^b Biostatistics and Bioinformatics Facility, JG Brown Cancer Center, University of Louisville, Louisville, KY, USA

^c The Christina Lee Brown Envirome Institute, University of Louisville, Louisville, KY, USA

ARTICLE INFO

Keywords:

Adaptive design
Adaptive signature design
Biomarker
Oncology
Precision medicine
Randomized clinical trials

ABSTRACT

Genomics having a profound impact on oncology drug development necessitates the use of genomic signatures for therapeutic strategy and emerging medicine proposals. Since its advent in the arena of clinical trials biomarker-related predictive methods for the identification and selection of patient subgroups, with optimal treatment response, are widely used. Genetic signatures which are accountable for the differential response to treatments are experimentally recognizable and analytically validated in phase II stage of clinical trials. The availability of robust and validated biomarkers in phase III is limited. Hence, the development of a clinical trial design without the availability of biomarker identity for treatment-sensitive patients becomes indispensable. Adaptive Signature Design (ASD) is a design procedure of developing and validating a predictive classifier (diagnostic testing strategy) when the signature of subjects responding differentially to treatment is remote in the context of the study. This review provides a detailed methodology and statistical background of this pioneering design developed by Freidlin and Simon (2005). In addition, it concentrates on the advances in ASD regarding statistical issues such as predictive assay identification, classification techniques, statistical methods, subgroup search, choice of differentially expressed genes, and multiplicity correction. The statistical methodology behind the design is explained with the intent of building the ground steps for future research approachable, especially for beginning researchers. Most of the existing research articles give a microcosmic view of the design and lack in describing the details behind the methodology. This study covers those details and marks the novelty of our research.

1. Introduction

The diversity and dynamic nature of cancerous tumors due to DNA alterations, and oncogenic mutations create their uniqueness [1]. Predictive biomarkers have the capability to distinguish patients influenced by treatment and are a great catalyst in the development of effective drugs. With the fast pacing research in the field of personalized medicine, genomic sequencing technologies and molecular targeted therapy, the inclusion of the personal entities need to be highly prioritized. This includes demographic factors such as age, sex, education, marital status, knowledge of biological influences, genetic aspects, and biomarkers. The dawn of gene-based approaches such as single nucleotide polymorphism (SNP) and the ground-breaking scientific advancement of the Human Genome Project (HGP) [2] modified the pathway of drug development and channelized it in a direction focusing on molecular and cellular variations in clinical response to therapy. The

traditional randomized clinical trials (RCTs) give importance to a singled-out query about the average treatment effect in the overall population, and thus this conventional approach needs personalization and re-designs. Presumption of homogeneity in the targeted population corresponding to treatment efficacy does not have an adequate scientific basis. Thus clinical trial designs consistent with modern heterogeneous tumor biology are destined to reshape the fundamentals of today's medicines, with drug targeted to disease [3]. Effective assessment of treatment effect requires the identification and validation of predictive biomarkers that can not only accurately measure the overall treatment effect but can assist in strongly inhibiting the target having the capability of recognizing the subgroup effect. Personalized medicine and single patient (N-of-1) trial [4] accurately determines the right drug to the right patient by accommodating individual needs, thereby saving time and resources. The ultimate goal of the development of molecularly targeted therapy is to improve the efficacy and selectivity

* Corresponding author. The Christina Lee Brown Envirome Institute, University of Louisville, Louisville, KY, USA.

E-mail addresses: arinjita.bhattacharyya@louisville.edu (A. Bhattacharyya), shesh.raai@louisville.edu (S.N. Rai).

of cancer treatment by exploiting the differences between cancer cells and normal cells [5]. In summary, the lack of a promising candidate signature before initiating phase III trials brings substantial complexity. To address this situation, adaptive signature design (ASD) [6] involving the partitioning of patients in training and validation sets for developing and confirming of a predictive classifier in a single (pivotal) trial, has been proposed. This article is a review of the research on adaptive designs in clinical trials and a comprehensive review of the research on ASD, along with its selected extensions. The review has a constructive flow with the sequence being: a) the role of adaptive design; b) definition of critical terms like biomarker; c) distinction between different types of biomarkers; d) biomarker-guided adaptive designs in phase III clinical trials; and e) statistical background of ASD; The discussion section focuses on recent advancements in ASD concentrating on gene-based classifiers, multiplicity issues, and subgroup effects.

2. Methods

A comprehensive review with keywords such as ‘adaptive signature design’, ‘precision medicine’, ‘biomarker-guided adaptive’, ‘subgroup effect in adaptive signature design’ identified the articles that have a direct contribution in extending and addressing the statistical issues in ASD by searching the internet mostly in Google Scholar database with a year restriction of 2005–2018. The remaining articles were mostly identified from the reference list of these previously included articles. The initial search was conducted by reading the abstract and discussion section of the articles. Those which mentioned ASD and its extensions, statistical aspects of the designs were studied thoroughly to extract information on the definition of the trial design(s), treatment group randomization, and design methodology, real-life clinical trial example of the design, leverage and limitations. We have considered 25 articles in the year 2016–2018, of which reference nos. 53, 54, 61, 62, and 66 (2017–2018) are the ones concentrated on ASD and its advancements described in Table 1. Also, these designs are explained in the discussion section of the article.

2.1. Role of adaptive design in clinical trials

We discuss the pre-specified and pre-planned modifications in the form of sample size, cost, and duration of the trial, the advantages, and the disadvantages of the prominent modern adaptive designs in clinical trials. The current draft of the Guidance for Industry provided by FDA defines adaptive design as the one that “includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study. Analyses of the accumulating study data are carried out at prospectively planned time points within the study, can be performed in a fully blinded manner or in an unblinded manner, and can occur with or without formal statistical hypothesis testing” [7]. Adaptive designs in phase II and phase III provide the flexibility and modulations on gathering analytically and clinically validated information, which can be in the form of patient inclusion/exclusion criteria, schedule of patient evaluations, clinical trial endpoints, and the number of interim

analyses, as well as the selection of treatment regimens. Adaptive designs are flexible to incorporate during-study modifications, such as a change in sample size, the primary endpoint, the null hypothesis like shifting from superiority to inferiority trial and primary to secondary endpoint [8], treatment re-allocation, stratification changes, as there might be limited knowledge regarding these decisions at the initial stage and requires ‘edit-as-study-progress’ and thus is efficient than traditional phase III designs like group sequential designs, where the trial stops as soon as there is sufficient evidence to reach a conclusion [9]. However, these modifications need pre-specified planning, and implementing these changes need to be done in a prospective manner to ensure data integrity. Adaptive clinical trials are more efficient in terms of time required and the number of participants. This modifications during the trials facilitate the likelihood of detecting treatment effects in clinical trials [10,11]. Table 2 gives a brief illustration of the prevalent types of adaptive designs and helps to understand the basic differences: Response Adaptive, Group Sequential, Sample-size Re-estimation, Seamless Design, and Biomarker-Adaptive [11–14].

Since the focus is on biomarker-guided adaptive designs and hence ASD, the definitions and demarcation of biomarkers have been accounted for in the next section.

2.2. What is a biomarker?

The Biomarkers Definitions Working Group defined a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”. Biomarkers contribute knowledge and provide guidance about clinical pharmacology and form the basis of designing clinical trials, with additional benefits of evaluating the safety and efficacy [24]. As an example, reduction of elevated arterial blood pressure have been used as a biomarker condition, which signifies the ability of antihypertensive drug depreciating the risk of nonfatal and fatal (total) stroke in isolated systolic hypertension, cognitive failure, etc. [25]. Biomarkers are one of the most critical features in enhancing the accuracy of diagnosis, prognosis assessment, and therapeutic targeting. The phase III clinical trial objective is to evaluate if the new treatment validated in phase II is beneficial to that targeted broader section of the patient population, and thus a comparison of outcomes between treatment and control group is undertaken to judge these benefits. A biomarker has the advantage of detecting the treatment effect in a subpopulation. However, the development and its validation involve complications in time and resources, so it lags the therapeutic development of the agent [26]. Biomarkers can be determined in numerous ways. For example, easily obtainable body fluids, such as plasma, serum, or urine, serve as a surrogate biological assay. Well known biomarkers of significant relevance are oestrogen receptor (ER), progesterone receptor (PR), and HER2/neu in breast cancer; BCR–ABL fusion protein in chronic myeloid leukemia (CML); c-KIT mutations in gastrointestinal stromal tumors (GIST); and epidermal growth factor receptor 1 (EGFR1) mutations in non-small cell lung cancer (NSCLC) [27]. The biomarkers are classified mostly into three categories: diagnostic, prognostic, and predictive.

Table 1
Summary of research on ASD (2017–2018).

Reference Nos. (2017–2018)	Details
53	Avoiding subgroup analysis when a lack of statistical significance for a subgroup is indicated by overall test, thus saving resources for future research
54	Novel methods for subgroup selection and estimation treatment impact.
61	Different classification algorithms with gene-main effect and gene-treatment interaction effect, predicting the efficacy of ASD.
62	early-phase proof-of-concept (POC) studies for cytotoxic oncology drugs such as single-arm adaptive signature design (ASD), single-arm Biomarker-adaptive threshold (BAT) design, designed for exploring the anti-tumor activity and to target biomarker-relevant topics, e.g., subgroup selection, biomarker threshold evaluation.
66	enhancement of ASD to improve robustness, and impact of allocation ratio between learning and confirm stage on the power of the design.

Table 2
Summary of adaptive designs.

Types of Adaptive Design	Main Concept & Adaptation	Examples
1. Response-Adaptive Design Advantages: <ul style="list-style-type: none"> ● Novel approach with increment in the accrual of patients leading to high success. Disadvantages: <ul style="list-style-type: none"> ● Possibility of incorrect decisions, and the introduction of bias. ● Statistical inefficiency due to an unequal number of patients in different treatment arms. ● Bias due to time-trends in the prognostic mixture of subjects. 	Alteration of the allocation ratio depending on the interim analysis.	BATTLE 2 ^a [15], ASTIN ^b [16,17]
2. Group-Sequential Design Advantages: <ul style="list-style-type: none"> ● Flexible approach, with room for sample size modification in a blinded manner, number, and spacing of interim analysis. Disadvantages: <ul style="list-style-type: none"> ● Parameter estimates, and confidence intervals are non-conventional. ● Data and safety monitoring demand attention for premature termination. 	Early stopping rule for safety, futility or efficacy.	DEVELOP-UK ^c trial [18], CAPTURE ^d trial [19,20]
3. Sample-Size Re-estimation Design Advantages: <ul style="list-style-type: none"> ● Small sample size upfront commitment; and adjustments due to unknown treatment effect, variance, lesser regulatory difficulties. Disadvantages: <ul style="list-style-type: none"> ● potential investigator-oriented bias introduction ● Interim treatment estimate can be misleading. 	The interim analysis determines the targeted sample size and may escalate or de-escalate.	DEVELOP-UK trial [18], NCT00103168 ^e [21]
4. Seamless Design Advantages: <ul style="list-style-type: none"> ● Information from dual phases. ● Time and cost saving; and no wait time for transition. Disadvantages: <ul style="list-style-type: none"> ● Elimination of the flexibility of modifying phase III based on phase II. ● Inadequate dose-response modeling leading to risk. ● Heterogeneity in data from two phases may cause issues. 	Sequential and immediate continuation from one phase to subsequent phase provided overlapping information.	NCT00543543 ^f [22]
5. Biomarker-Adaptive Design Advantages: <ul style="list-style-type: none"> ● Biomarker status helps in channelizing treatment towards biomarker positive (MK+) or biomarker negative (MK-) group; eligibility criteria modified depending on futility monitoring of MK- group. ● Identification of the natural course of a disease and achieving early detection of disease. Disadvantages: <ul style="list-style-type: none"> ● Demands clinical and analytical validity of biomarkers, statistical challenges, operational complexity. 	Adaption on the collection and validation of biomarker information.	FOCUS4 [23]

^a Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination.

^b Acute Stroke Therapy by Inhibition of Neutrophils.

^c Donor Ex Vivo Lung Perfusion in UK lung transplantation.

^d c7E3 Fab antiplatelet therapy in unstable refractory angina.

^e European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group Intergroup Randomized Trial.

^f 9-valent HPV vaccine trial against infection and intraepithelial Neoplasia.

- i) **Diagnostic Biomarkers:** Diagnostic Biomarkers help in the diagnosis of the disease and development of new therapeutics by monitoring disease progression in a population. The biomarker projects its benefits in disease assessment when the trial initiates. Testing of an individual with an internal, patient-specific control is more important than comparing to a “normal reference sample”. For example, plasma-based microRNA (miRNA) diagnostic assay was developed in the study for colorectal cancer (CRC) where seven miRNAs (miR-21, miR-29c, miR-122, miR-192, miR-346, miR-372, and miR-374a) were selected based upon p-value, depending on ANOVA or *t*-test as specified in the design phase for multiple-comparisons, or two-way. Area under the curve (AUC), fold change, and biological plausibility produced on an average 77% accuracy in the validation cohort for the respective comparisons [28]. Plasma NfL and CSF biomarkers have a good diagnostic performance to detect Alzheimer's disease in adults with Down syndrome [29].
- ii) **Prognostic Biomarkers:** Prognostic Biomarkers are used to detect disease recurrence or progression in patients who have the disease or medical condition of interest, regardless of the therapy which helps in the selection of patient groups for a specific treatment. For example, biomarkers providing information about a relapse of disease in patients undergoing treatment and those with progression-free survival in patients with metastatic disease. A prognostic biomarker is also a baseline patient characteristic independent of

- therapy or treatment which categorizes patients on the basis of the degree of the outcome [30]. Studies to identify prognostic gene signatures such as sentinel lymph node (SLN) to aid in risk stratification of patients with melanoma tumor-positive SLNs, are carried out with the application of Quantitative reverse-transcriptase polymerase chain reaction (RT-PCR) assay for validation of the gene expression with de-identified human subject data from the Sunbelt Melanoma Trial (SMT), involving enrollment of about 3600 patients over 79 centers throughout North America in a randomized, prospective trial, with disease-free survival (DFS) as the primary endpoint. Two SLN genes (PIGR and TFAP2A) provided high prognostic accuracy, i.e. these biomarkers can detect the progression of the disease regardless of the treatment in SLN positive melanoma patients, with the area under the receiver operating characteristic curve (AUC) technique, along with comparing it with the current American Joint Committee on Cancer (AJCC) staging system. This study was conducted with a goal of contributing towards individualized patient risk measurement, and this SLN gene signature, when combined with clinico pathological features, serves aptly the purpose of better guiding treatment decisions in the near future [31].
- iii) **Predictive Biomarkers:** Predictive Biomarkers provides information about the favorable and unfavorable effect of therapeutic intervention. A predictive biomarker is a baseline characteristic

which categorizes patients by their degree of response to treatment. A predictive biomarker is utilized to detect the treatment effect between a biomarker-positive (MK+) and a biomarker negative (MK-) group. However, the use of the predictive biomarkers in phase III poses challenges, when the inclusion of MK- patients is halted in the absence of strong biological evidence of treatment effectiveness in that group [32]. For example, certain cystic fibrosis transmembrane conductance regulator (CFTR) mutations may be used as predictive biomarkers in clinical trials for treatment response evaluation for cystic fibrosis [33].

There is a subtle difference between prognostic and predictive biomarker; a prognostic biomarker detects the path of disease irrespective of treatment, and predictive biomarker determines the response to treatment when the disease is already present. Biomarkers may have both prognostic and predictive characteristics. For example, the estrogen receptor (ER) can be prognostic in breast cancer, where patients diagnosed with ER-negative have a higher risk of relapse than ER-positive patients with a similar disease stage. Also, ER can be predictive of benefit from tamoxifen [34], where antiestrogen tamoxifen is more effective in preventing breast cancer recurrences in ER-positive patients than in ER-negative patients.

2.3. Biomarker guided adaptive Phase-III clinical trials

An encompassing review of biomarker-guided trial designs has been provided in these two articles, one focusing on adaptive trial designs [35] and the other on non-adaptive trial designs [36]. One can also refer to their website for a detailed user-friendly approach where each of the designs is explained in a graphical manner, with pros and cons [37]. Adaptive designs target all patients with the plan of treatment efficacy determination in all of them, if not significant initiates' subgroup finding with treatment benefits. In settings where a single continuous candidate biomarker is available, but its positivity threshold is not predefined, adaptive threshold plans try to find such a threshold. Adaptive biomarker designs have been proposed to evaluate multiple binary biomarkers defined in advance. We are mostly dedicated to describing ASD among the many biomarker-guided designs. ASD comes in six identical and interchangeable versions [38]:

i) General Adaptive Signature Design; ii) Adaptive Signature Design; iii) Adaptive Threshold Design, iv) Cross-validated Adaptive Signature Design; v) Molecular Signature Design; vi) Adaptive Signature Design with Subgroup Plots. We have also focused on giving a detailed overview of the statistical integrities of the logistic regression model used in ASD.

In most of the biomarker dependent adaptive designs, it is assumed that we know the number and identity of differentially expressed genes. However, in reality in a large pool of evaluated genes, the significant ones need to be determined initially by high throughput screening methods, and then we can proceed to the estimation of gene-treatment interaction effect like in ASD. Also, the choice of significance level (α) for differentially expressed gene selection by Bonferroni adjustment and False Discovery Rate (FDR) need to be assessed by our proposed design Unknown Genes Adaptive Signature Design (UKG-ASD). A schematic diagram of Known Genes Adaptive Signature Design (KG-ASD) and (UKG-ASD) is given in Fig. 1.

This section depicts the Adaptive Signature Design (ASD) and its major role as a biomarker-guided design. The design and the statistical formulations behind it are well-defined here along with its equivalent extensions with the purpose of increment in efficiency and statistical power.

2.4. General Adaptive Signature Design

In this design, the candidate biomarkers are selected from a large set of biomarkers, and the threshold point is optimized using a training set,

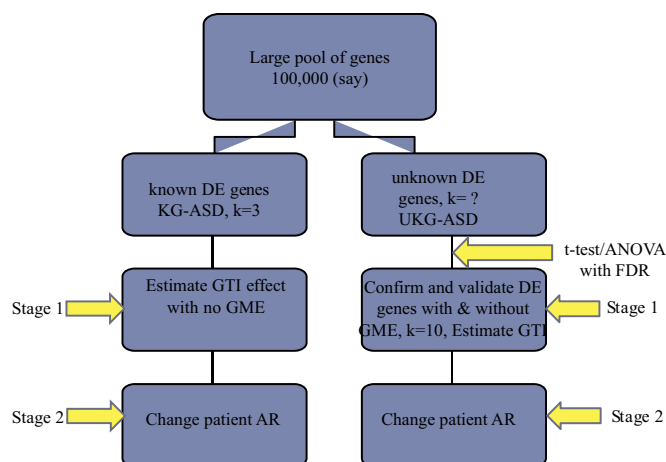


Fig. 1. Summary of Known Gene Adaptive Signature Design (KG-ASD) and, Unknown Gene Adaptive Signature Design (UKG-ASD)

AR: Allocation Ratio; FDR: False-Discovery Rate; DE: Differentially Expressed; GME: Gene Main Effect; GTI: Gene Treatment Interaction.

and the assessment of the signatures is carried out in the validation set [38]. Power restriction in the biomarker-defined subset imposes a major setback for the design. This approach finds its use when there are a number of available candidate biomarkers, but data from a Phase III setting is required for choosing the most appropriate biomarkers [39].

2.5. Adaptive signature design (ASD)

ASD was proposed by Freidlin and Simon [6], also known as Biomarker Adaptive Signature Design, Two-stage ASD, or Adaptive Two-stage Design is a design for randomized clinical trials of targeted agents in settings where an assay or signature that identifies sensitive patients is not available at the outset of the study. The trial, upon having no indications for overall treatment effect, proceeds to divide the patient population in two mutually exclusive groups. One group is used to set up a parameter to identify the subset of patients where the treatment is most likely to be effective – the parameter designated as “indication classifier”. The other group of patients is used to see the efficacy of the treatment in that particular subset [40]. The fundamental structure of the design relies on subgrouping the population into two complementary cohorts, after the completion of the trial. Potential biomarkers are selected and treatment efficacy is investigated in the overall population and the biomarker-qualified subpopulation within the validation cohort. Upon validating and standardizing the biomarker assay, ASD is thoroughly designed and crafted in a statistically enriched way so that it preserves the type-I error rate, simultaneously identifying an optimally targeted design [41]. Recent application of ASD includes evaluation of efficacy of the MAGE-A3 immunotherapeutic in patients with stage IIIB or IIIC melanoma in the adjuvant setting with disease-free survival as the outcome. Median disease-free survival was 11.0 months (95% CI 10.0–11.9) in the MAGE-A3 group and 11.2 months (8.6–14.1) in the placebo group (hazard ratio [HR] 1.01, 0.88–1.17, $p = 0.86$). However, treatment-related adverse events of grade 3 or worse were reported within a month from the start of the treatment barring the development of the MAGE-A3 immunotherapeutic for use in melanoma [42]. The design is explained in a graphical format in Fig. 2.

The three main steps of the design:

- i) Statistically valid identification, dependent on the first stage of the trial, of the subset of patients who are most likely to benefit from the new agent;
- ii) Properly powered test of overall treatment effect at the end of the trial with all randomized patients; and
- iii) Test of treatment effect for the subset identified in the first stage but

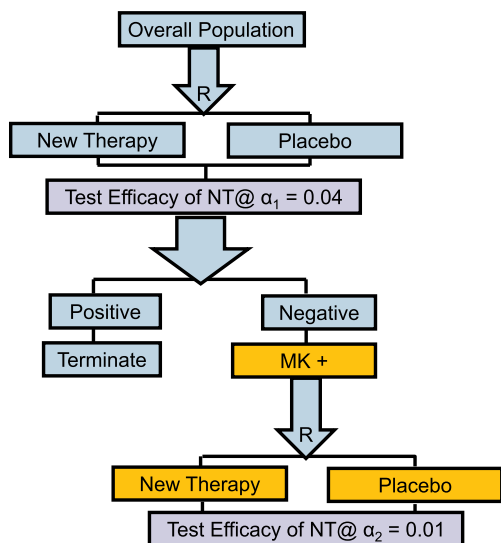


Fig. 2. Adaptive signature design (ASD) R: Randomize; NT: New therapy; MK⁺: Biomarker positive.

only with patients randomized in the remainder of the trial [6].

The results of the trial influence the primary plan for the final analysis, so the design is adaptive and is at par with the definition provided by the FDA Guidance draft.

2.6. Statistics behind Adaptive Signature Design (ASD)

- i) There is an accrual of N patients who are subjected to the new proposed treatment to determine its effect, i.e. outcomes for all patients randomized to receive the proposed new drug is compared with outcomes for all patients randomized to receive placebo. The overall test of treatment effect over control is done by the Pearson Chi squared test when the response variable is binary and declares that the test is significant at a two-sided significance level $\alpha_1 = 0.04$ (say). This methodology though explained in case of binomial outcomes, with minute modifications, have its applicability to time-to-event end points, such as survival or disease-free survival.
- ii) If the overall test is found non-significant, a subset analysis approach is undertaken where the patient set is randomly partitioned into a training set and a validation set consisting of n_1 and n_2 patients, $n_1 + n_2 = N$, and conducting the signature development and its validation in a two-stage process. Here, the randomization ratio is 1:1, simple unbiased randomization. It is obvious that the subjects with overexpressed sensitive genes will respond to treatment more than the subjects with non-sensitive genes. This indicator of sensitivity is the biomarker information that is unknown. It is assumed that there are L evaluated genes out of which k genes are assumed to be sensitive; however, the identities of them are unknown. Dimension reduction method such as principal component analysis, partial least squares, and LASSO can also be used to select these k differentially expressed genes. The scenario of the presence of numerically large candidate genes can summon various variable selection methods to include only variables that can distinguish outcome on treatment arm from that of the control arm. However, it was suggested that accurate classification of patients should be the foremost priority instead of the statistical significance of the individual variables [41]. The methodical approach to classify patients as sensitive/non-sensitive based on this unknown hidden label, is to develop an “indicative classifier” that identifies the subset of patients who are profitable from the new treatment compared to control out of these n_1 patients in the training set,

where the treatment-gene interaction estimates for each single gene corresponding to each individual is tested to see if they are significant at level $\eta = 0.025$ (say). Let m genes be significant out of these k sensitive genes.

- iii) Finally, with these m genes, the model is validated with the remaining n_2 patients, since the training set of n_1 patients have been utilized to build the classifier. The odd's ratios with gene-expression data from the validation set and estimated coefficients from the training set are calculated and compared to see if they exceed a pre-specified threshold R , for at least G genes out of these m sensitive genes, for every subject which can eventually declare that individual to be sensitive.
- iv) To investigate the effect of changing the randomization ratio in a group of patients who are more sensitive to the treatment arm over the control arm, the subset treatment effect test is conducted at the reduced significance level of $\alpha_2 = 0.05 - \alpha_1 = 0.01$. The type-I error rate can be partitioned, and the split ratio can be optimized to minimize the total sample size, constrained on the statistical power of the overall analysis and subset analysis.

v) The power is calculated $\Phi \left\{ \frac{(P_t - P_c) - z_{1-\alpha} \left(\frac{p^-(1-p^-)^4}{N} \right)^{0.5}}{\left[\frac{P_t(1-P_t)^2}{N} + \frac{P_c(1-P_c)^2}{N} \right]^{0.5}} \right\}$, $p^- = \frac{P_t + P_c}{2}$

Where, P_t : the expected probability of response in treatment & P_c : the expected probability of response in control where, $\Phi ()$ is the normal probability function, and $Z_{1-\alpha}$ is the

(1- α) th percentile of the normal distribution.

For simplicity, we are explaining the statistical model, when the response is binary.

$y_i \sim Ber(1, p_i)$, where y_i is the outcome variable, $y_i = 1$ with probability p_i and

$E(y_i|X) = p_i$, $i = 1, \dots, N$, X is the matrix of covariates consisting of fixed treatment effect, gene main effects and gene-treatment interaction effects, z_{ij} is the gene expression value corresponding to i th individual and j th gene. $y_i = \mu_i + \varepsilon_i$, where, μ_i is the mean effect and ε_i is the error. $p_i = \frac{\exp(B'X)}{\exp(B'X) + 1}$, where B is set the coefficients, N is the total no. of patients, β_0 is the intercept term, β_1 is the treatment main effect. Thus,

$$\begin{aligned} \text{logit}(p_i) &= \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 t_i + \delta_1 z_{i1} + \dots + \delta_m z_{im} + \delta_{m+1} z_{i(m+1)} \\ &+ \dots + \delta_k z_{ik} + \gamma_1 t_i z_{i1} + \dots + \gamma_m t_i z_{im} + \dots \\ &+ \gamma_{m+1} t_i z_{i(m+1)} + \dots + \gamma_k t_i z_{ik}, \quad i = 1 \dots N \end{aligned}$$

The probability of response is p and the treatment status is t_i , where $t_i \in \{0,1\}$. The gene main effects of the sensitivity and non-sensitivity genes ($\delta_1, \delta_2, \dots, \delta_k$) and the interaction effects of the non-sensitivity genes ($\gamma_{m+1}, \dots, \gamma_k$) are assumed to be zero. So, the model reduces to

$$\text{logit}(p_i) = \beta_0 + \beta_1 t_i + \gamma_1 t_i z_{i1} + \dots + \gamma_m t_i z_{im}, \quad i = 1 \dots N$$

In stage 1, with the training set for each i^{th} individual single gene logistic model is fitted, which implies fitting m models for each of the n_1 individuals, i.e.

$$\text{logit}(p_i) = \beta_0 + \beta_1 t_i + \gamma_1 t_i z_{i1}, \dots, \text{logit}(p_i) = \beta_0 + \beta_1 t_i + \gamma_m t_i z_{im}$$

Classification development in this design is not restricted to the covariates considering gene expression measurements, and hence can be extended with a proportional hazard model in survival data analysis, where the algorithm of classifier development is being described in detail considering overall survival as the primary endpoint for new drug approval of castration-resistant prostate cancer (CRPC) [41]. As power is a monotonically increasing function of sample size, the simple concept of signature development and validation restricts the power in cases where large sample size is required for signature development in high dimensional data and enough screening needs to be done to

Table 3
Summary of adaptive design using biomarkers.

Types of ASD	Modifications	Examples
1. Adaptive-Signature Design (ASD)	Identification and validation of biomarker and utilizing the information to propose an MK + subset.	MAGE-A3-DERMA trial [42]
<p>Advantages:</p> <ul style="list-style-type: none"> • Distinguishing and validating of genomic signature in a single trial. • Finding the optimal group of patients benefiting from the proposed new treatment. • The significance level is preserved without inflating it as the design takes a split-sample approach and uses a different sample set of patients for predictive biomarker identification and validation. • A novel methodology, with no weights for randomization ratio, without statistical adjustment, making the design-unbiased, increasing the efficiency and the expediting the treatment approval. <p>Disadvantages:</p> <ul style="list-style-type: none"> • Minimal difference between MK+ and MK- patient group demands a large sample size. • Since only the validation group of patients is utilized for biomarker validation, due to the split-sample technique, achieving the desired power for the design is somewhat restricted. • The trial must come to closure before the treatment comparison can be carried out. 		
2. Adaptive Threshold Design	The treatment effect comparison is tested over the broad population with establishing and validating a cut-off point for a pre-specified signature assay, so that the levels of the continuous biomarker can be converted into a dichotomous (positive/negative) variable, thus, detecting treatment sensitive subpopulation.	VACURG ^a trial [48]
<p>Advantages:</p> <ul style="list-style-type: none"> • Pre-specified cut off point is not essential for candidate biomarker validation increasing efficiency. • Identification of an optimal cut-off point for detecting sensitive patients (i.e., biomarker-positive patients). • Moderating dependency on Phase II data for establishing test cut-off point. <p>Disadvantages:</p> <ul style="list-style-type: none"> • The requirement of a quantitative biomarker before the design is applied. • Power analysis and treatment comparison are done at the end of the trial. • Accrual of patients is biomarker status independent. • Inadequate power if the treatment effect is restricted to a relatively small fraction of the study population. • Larger sample size may be required and can cause power redundancy. Data from the same study for defining and validating cut-off point may generate bias. 		
3. Cross-validated Adaptive Signature Design (CV-ASD)	Like ASD, sensitive patient classifier development and validating with a cross-validation method ignored.	EORTC 10,994 [49]
<p>Advantages:</p> <ul style="list-style-type: none"> • Enhances the performance of ASD with the increase in power and efficiency, permitting the maximization of the portion of study patients contributing to the development of the diagnostic signature. • Maximizes the sensitive subgroup size. <p>Disadvantages:</p> <ul style="list-style-type: none"> • Multiplicity problem and similar challenges as ASD. 		
4. Molecular Signature Design	A similar approach to that of ASD with a survival endpoint.	No real-life example found
<p>Advantages:</p> <ul style="list-style-type: none"> • Considered as a promising strategy for drug development as it takes advantage of the use of an end-point with clear clinical gain. <p>Disadvantages:</p> <ul style="list-style-type: none"> • No information. 		
5. General Adaptive Signature Design	A similar approach of ASD to identify and validate the biomarkers chosen.	No real-life example found
<p>Advantages:</p> <ul style="list-style-type: none"> • Optimizes the test based on randomized data for patients in the Phase III setting. <p>Disadvantages:</p> <ul style="list-style-type: none"> • Power restriction on the MK + subgroup. 		

^a Veterans Administration Cooperative Urologic Research Group.

increase the proportion of sensitive patients reducing efficiency.

Following the guidelines of Antoniou et al. [35], a short overview of the five equivalent labels of ASD are described and presented in Table 3.

2.7. Adaptive Threshold Design

This is a type of design in which a putative biomarker is measured on a continuous or graded scale. The treatment effect comparison is tested over the broad population by establishing and validating a cut-off point for a pre-specified signature assay. Thus, the levels of the continuous biomarker can be converted into a dichotomous (positive/negative) variable and enables detecting a treatment sensitive subpopulation. This design first tests the treatment effect in the overall population, whose statistical significance claims the treatment to be effective in the broad eligibility population. Otherwise, this design provides a statistically rigorous assessment of the presence of a cut-off point that can identify a subgroup of patients that would benefit from the treatment without relying upon any arbitrary pre-specified cut-off

value. ASD has the objective of developing and validating a biomarker. On the other hand, this variant tries to identify and validate an optimal threshold point for a pre-specified biomarker, and converts the signature value into a binary classifier along with providing confidence interval of the cut-off point [40,43,44]. With no significant treatment effect in the broad population during, the entire population is divided into two parts. One half is utilized in developing a biomarker signature that can be predictive and obtain the sensitive population, and the other half is used in validating the claim. This design focusses on the establishment and validation of a cut-off points/threshold for a pre-specified biomarker for characterization of the sensitive population subset, upon finding no treatment effect in the entire population. One of the major differences between this design and ASD is the utility of human samples for measuring a pre-specified biomarker. However, it is not targeted as the eligibility criteria for patient inclusion. This design has two variations in conducting the analysis:

2.7.1. Plan A

It is a three-step process of confirming: 1) the benefits of experiment

treatment arm over placebo in the overall population; 2) subset of patients having biomarker values greater than a threshold c (say) are prone to greater effectiveness in the treatment arm than the control; and 3) no detection of treatment effect. This method is at par with ASD with overall and subset test being carried out at a significance level of $\alpha_1 = 0.04$ (say), $\alpha_2 = 0.01$ (say); $\alpha_1 + \alpha_2 = \alpha = 0.05$ (say). Using all possible cutoff values, the test statistic is calculated as the maximum of the log-likelihood ratio statistic for treatment effects having biomarker values above the cutoff value.

2.7.2. Plan B

This is considered as a generalization of Plan A where the differential treatment effect in the population is detected in the first stage. The second stage devotes in obtaining the biomarker threshold above which the treatment effect demonstrates beneficial acts towards the patient population. In this method, test statistic using the maximum overall cut-off values as in Plan A is recommended. The efficiency is administrated by indulging the correlation structure of the test statistics in combining the overall and subset tests, with the test statistic as the larger of the test statistic for Plan A and the log-likelihood ratio statistic for treatment effect in the entire population. A bootstrap re-sampling approach was undertaken for the point and interval estimates of the cut-off value, as Plan B being more productive than Plan A but with slightly larger sample size and redundant power.

2.8. Cross-validated adaptive signature design (CV-ASD)

This design was proposed to increase the power and efficiency of ASD. The approach is similar to that of ASD with the objective of identifying the genomic signatures in the training set, and validating in the test set if the overall treatment effect is not significant [45]. The main factor that differentiates CV-ASD from ASD is the use of a K-fold cross-validation procedure, so that the entire study population is utilized in both signature development and validation steps with the superiority monitoring in the MK + group. The accrued population after the completion of the trial is divided into two non-overlapping cohorts: a predictive signature development cohort and a validation cohort, having N_1 and N_2 patients respectively. With $K = 10$ (say) cross-validation, the development cohort has 10 divisions each with $N_1/10$ patients forming d_1, d_2, \dots, d_{10} groups. Similarly, v_1, v_2, \dots, v_{10} groups are formed out of the validation cohort each with $N_2/10$ patients. With each $d_k, k = 1, 2, \dots, 10$, a predictive signature is developed whose application entails the identification of s_k , a subset of sensitive patients out of each v_k set. This process has a major contribution in utilizing all the patients, and the union of all the sensitive set of patients forms the total sensitive population. The design enhances the performance of ASD in terms of maximizing the population engagement in signature development and justification and increasing the power.

2.9. Molecular Signature Design

This is a Phase III design with the motive of comparing the new drug with a standard of care and is on the similar lines as of ASD, except that the primary endpoint is overall survival instead of binary [46]. The accumulation of tissue samples from patients is done at baseline. However, the analysis of them is expected to be performed at the near end of the trial, when all possible biomarker combinations are utilized to propose a classifier that can distinguish the patients sensitive to the new regimen with overall survival as the outcome. The trial is designed as follows:

- i) Collect biomarker tissue samples on all individuals when the trial begins.
- ii) Randomize patients to new therapy versus placebo, and if the effect of treatment is significant at the level 0.04 with the survival outcome, the trial stops.

- iii) If the test is not statistically significant, combinations of gene signatures are utilized to form a biomarker-based classifier to identify subjects sensitive to new care in the training set with a certain proportion of patients, and the performance of the classifier is determined in the validation set with the remaining proportion of patients.

2.10. Adaptive signature design with subgroup plots

Adaptive Signature design with Subgroup Plots [47] is an extension of ASD, which brings flexibility and uses a tail-oriented or sliding window subgroup plots for subset identification by assessing several cut-off points of the benefit score obtained by the subgroup plots. The subsets consist of patients responsive to the treatment. The benefit threshold is defined as the difference in effect between the new and the standard mode of care. In this way, it provides broader confidence intervals of the estimated treatment benefit. No statistical consideration has been found for this approach. The method is explained by generating a hypothetical dataset of 10,000 markers from a normal distribution and 400 patients, however, no real-life example has been found. Expression levels for each marker were generated under independent normal distributions. For each of the treatment and control groups, the binary outcome is generated, and can be extended with survival outcome. The estimated treatment benefit and its confidence intervals showed that they are within the pre-specified benefit threshold of experiment vs control, and thus can proceed for subgroup analysis. 99% confidence intervals and confidence bands are constructed, which makes up part of the sliding window and tail-oriented subgroup plots. A subgroup is identified when it has a lower bound of simultaneous 99% confidence intervals greater than the benefit threshold.

3. Discussion

Cancer research is reaching a new epitome with the advancement of implementing biomarkers in accurate diagnosis, prevention, and therapeutic treatment of diseases. One of the key landmarks in the exponential rise in research of melanoma studies is the precision medicine, which considers variability in genes, environment, and lifestyle per individual. The use of predictive biomarkers for the identification of the patient population benefiting from the treatment makes the drug development process complicated and expensive. However, it is scientifically and fundamentally authenticated that with an increase in the probability of success rate of oncology drugs, reducing the number of patients exposed finally to these expensive drugs, and avoiding adverse effects can curb the public health care expenditure [50].

Among the several challenges in biomarker validation, the following areas were identified which need attention in ASD and in general biomarker adaptive designs.

Subgroup Finding: Subgroup analysis is referred to as an evaluation of treatment effects in specific subgroups of patients defined by baseline characteristics, where there is heterogeneity of treatment effects across a subset of patients. Detection of subgroup effect helps in directing patients to the treatment arm, as they will be profitable from treatment exposure. An interaction test between the treatment and subgroup is a commonly used statistical method for assessing the heterogeneity of treatment effects among subgroups of a baseline (predictive) variable. Also, the prevalence of minimal sensitive genes can lead to miscalculation of the treatment and subgroup effect [51].

Recent researches with ASD include extensions in the subgroup analysis arena where inclusion of futility or decision rule can contribute to the prevention of cost and complications of defining subgroups based on complex and expensive biomarkers, such as multivariate Quantitative polymerase chain reaction (qPCR). Frequentist and Bayesian approaches based on conditional power and predictive power respectively, with continuous efficacy endpoint restricts conduction of

subgroup analysis, if the overall test confirms lack of statistical significance in the subgroup [52]. New techniques of subgroup selection based on utility function whose formulation includes the subgroup size and clinical indicator with the objective of maximizing power for treatment effect in the selected subgroup with baseline covariates such as age, gender, systolic blood pressure, heart rate, a simple risk index and binary endpoints is addressed. There is no overall treatment effect evaluation as the test with the utility function is sufficient for the selected subpopulation cohort which could be the entire population as well [53]. Specific inclusion criteria defining the boundaries of clinical trial designs, with an eye on the comparison of the size of treatment effect inside and outside the subpopulation, and a balance between the size of the treatment effect and the subpopulation should be considered in defining the adaptive designs with subgroup selection in mind [54].

Choice of Gene Signatures: The choice of differentially expressed gene signatures before the ASD trial and the effect of it on the empirical power of the design need to be evaluated. There may be a single or no significant gene signature selected by the variable selection methods like PCA or LASSO, though the main purpose is the classification of patients in a treatment benefiting and non-benefiting groups. The process of identification of diagnostic [28] and prognostic [31] biomarkers is being addressed. The outlined objective is to predict the specific therapeutic outcomes, and the presence of complicated structure of thousands of genes in an experiment necessitates the selection of a subgroup of differentially expressed genes, which can be done by an appropriate test statistic like t -, U-(Mann-Whitney), M-statistics and calculating the p -values for gene ranking. Family-wise error and false discovery rate (FDR), false-positive error-controlled procedures, and receiver-operating characteristic (ROC) approach help in identifying significant number of genes [55].

Multiplicity Correction: The level of significance for multiple hypothesis testing can be adjusted by methods such as Bonferroni adjustment and FDR [56]. Correction of multiplicity technique, by finding the risk of false-positive findings in case of different scenarios of CVASD is considered. This addresses the drawbacks of biomarker-stratified and biomarker-enriched designs in which more than one potential predictive biomarker is being proposed. Inflation of type-I error is a major issue in “testing-in-all-direction” approach (TIADA) and CV-ASD. Accordingly, tweaking the pre-specified significance level with recalibration can be an appropriate substitute to handle the multiplicity issue which arises due to the inclusion of several biomarkers [57].

Classifier Development: It is well claimed that cancer therapies produce significant effects in a subgroup of patients, and hence the requirement of a well-validated predictive classifier with accurate reproducible predictions is inevitable. Performance of a classifier depends on the prevalence proportion, sample size, and the identified biomarker set. There can be multiple biomarker classifiers which can produce a similar effect in classification between sensitive and non-sensitive patients, and choosing the ideal and optimal one becomes difficult. A detailed review of the parametric [58] and nonparametric [59] methods of classification and dimension reduction of clinical outcomes using high-throughput informatics can form a basis for applications in melanoma studies. These classification algorithms can be applied in ASD. The extension of ASD with scenarios of solely gene expression main effect, treatment main effect, gene-treatment main effect, gene expression-treatment interaction effect using 10-fold cross-validation for dividing the data into training and testing sets is undertaken. The odd's ratio technique of classification is replaced by: i) Weighted Voting Kernel density analysis (KDA); ii) Weighted Voting Quadratic discriminant analysis (QDA); and iii) Weighted Voting Linear Discriminant Analysis (LDA) methods. The empirical power for QDA outperforms other methods [60].

Other research areas include developing ASD for single arm studies in phase I. For scenarios having strong biological reasoning and pre-clinical evidence to bolster the fact that molecular targeted therapy (MTA) may be differentially beneficial to a general population, early-

phase proof-of-concept (POC) studies for cytotoxic oncology drugs such as single-arm adaptive signature design (ASD), single-arm Biomarker-adaptive threshold (BAT) design, and Maximum Test Statistics approach can be designed for exploring the anti-tumor activity, and to target biomarker-relevant topics, e.g., subgroup selection, biomarker threshold evaluation [61].

Simon and Wang [62] focus on the development and use of classifiers based on biomarkers in tumors in oncology. Conditional bias correction problem was given in case of non-significant overall treatment effect, and no pre-defined assay based on polynomial functions were considered in treatment effect estimation in the MK + group [63]. This concept was however argued as the marker is constructed using the training sample, and a potential chance of correlation between log hazard ratio estimator on the training sample and the log hazard ratio estimator on the MK + subgroup is probable. Thus concentrating on the unconditional bias of estimators of the treatment effect for MK + patients is suggested [64]. Two important decisions in ASD i.e. the split of the significant level and the optimal allocation of patients in learn and confirm stage, and their impact on the power, supported by extensive simulations are also being addressed. In oncology drug development, there are set of biomarkers which modify the path of drug development as well as influence the prediction of biomarkers, resulting in meaningful patient sensitivity prediction. However, many a times, the availability of patients and biomarkers are limited. This leads to the need for conducting ASD. The design is implemented with a Cox-Proportional Hazard model and is available in an open-source R package “simASD” [65]. Relevant statistical aspects of interaction test for predictive biomarker identification, sample size, and power, validation of the classification model, etc. are described with ASD as a supplementary test when the test for overall treatment effect is not significant [66]. Future research scope includes: study of the effect of the choice of gene signatures on the power of ASD; incorporation of multiple gene effects by multigene penalized logistic regression models; study of effects of violation of the assumption of the normality of the genes; application to real data scenarios; and proper identification of the tuning parameters R and G . Thus, this study along with the recent advancement in ASD will help to extend the research in this arena.

Disclosure of potential conflicts of interest

Authors do not have any conflict.

Acknowledgments

This work was supported by the National Institute of Health grant P42 ES023716 to principal investigator: Dr S Srivastava and the National Institute of Health grant 1P20 GM113226 to principal investigator: Dr C McClain. Dr. Rai was also partially supported by Wendell Cherry Chair in Clinical Trial Research.

References

- [1] R.M. Simon, Genomic Clinical Trials and Predictive Medicine, Cambridge University Press, Cambridge, 2013, <https://doi.org/10.1017/CBO9781139026451>.
- [2] F. Carrasco-Ramiro, R. Peiró-Pastor, B. Aguado, Human Genomics Projects and Precision Medicine, (2017), <https://doi.org/10.1038/gt.2017.77>.
- [3] S. Matsui, M.E. Buyse, R.M. Simon, Design and Analysis of Clinical Trials for Predictive Medicine, n.d.
- [4] W.A. Demeyin, J. Frost, O.C. Ukoumunne, S. Briscoe, N. Britten, N of 1 Trials and the Optimal Individualisation of Drug Treatments: a Systematic Review Protocol, (n.d.). doi:10.1186/s13643-017-0479-6.
- [5] F.M. Balis, Evolution of anticancer drug discovery and the role of cell-based screening, J. Natl. Cancer Inst. (2002), <https://doi.org/10.1093/jnci/94.2.78>.
- [6] B. Freidlin, R. Simon, Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients, Clin. Cancer Res. (2005), <https://doi.org/10.1158/1078-0432.CCR-05-0605>.
- [7] Fda, Adaptive design clinical trials for drugs and biologics, Draft Guid (2010), <https://doi.org/10.1186/1477-7525-4-79>.
- [8] G. Hommel, Adaptive modifications of hypotheses after an interim analysis, Biom. J. 43 (2001) 581–589, [https://doi.org/10.1002/1521-4036\(200109\)43:5<581::AID-BIOM581>3.0.CO;2-1](https://doi.org/10.1002/1521-4036(200109)43:5<581::AID-BIOM581>3.0.CO;2-1)

- 43:5 < 581::AID-BIMJ581 > 3.0 CO;2-J.
- [9] C. Jennison, B.W. Turnbull, *Biometrika Trust Adaptive and Nonadaptive Group Sequential Tests*, (2006) https://www.jstor.org/stable/20441256?seq=1&cid=pdf-reference#references_tab_contents, Accessed date: 12 October 2018Accessed.
- [10] J.J.H. Park, K. Thorlund, E.J. Mills, Critical concepts in adaptive clinical trials, *Clin. Epidemiol.* (2018), <https://doi.org/10.2147/CLEP.S156708>.
- [11] D.A. Berry, Adaptive clinical trials in oncology, *Nat. Rev. Clin. Oncol.* (2011), <https://doi.org/10.1038/nrclinonc.2011.165>.
- [12] D.L. Bhatt, C. Mehta, Adaptive designs for clinical trials, *N. Engl. J. Med.* 375 (2016) 65–74, <https://doi.org/10.1056/NEJMr1510061>.
- [13] P. Pallmann, A.W. Bedding, B. Choodari-Oskoei, M. Dimairo, L. Flight, L.V. Hampson, J. Holmes, A.P. Mander, L.L. Odoni, M.R. Sydes, S.S. Villar, J.M.S.S. Wason, C.J. Weir, G.M. Wheeler, C. Yap, T. Jaki, Adaptive Designs in Clinical Trials: Why Use Them, and How to Run and Report Them, *BioMed Central*, 2018, <https://doi.org/10.1186/s12916-018-1017-7>.
- [14] E.L. Korn, B. Freidlin, Adaptive clinical trials: advantages and disadvantages of various adaptive design elements, *J. Natl. Cancer Inst.* 109 (2017), <https://doi.org/10.1093/jnci/djx013>.
- [15] V. Papadimitrakopoulou, J.J. Lee, I.I. Wistuba, A.S. Tsao, F. V Fossella, N. Kalthor, S. Gupta, L.A. Byers, J.G. Izzo, S.N. Gettinger, S.B. Goldberg, X. Tang, V.A. Miller, F. Skoulidis, D.L. Gibbons, L. Shen, C. Wei, L. Diao, S.A. Peng, J. Wang, A.L. Tam, K.R. Coombes, J.S. Koo, D.J. Mauro, E.H. Rubin, J. V Heymach, W.K. Hong, R.S. Herbst, The BATTLE-2 study: a biomarker-integrated targeted therapy study in previously treated patients with advanced non-small-cell lung cancer, *J. Clin. Oncol.* 34 (2016) 3638–3647, <https://doi.org/10.1200/JCO.2015.66.0084>.
- [16] A.P. Grieve, M. Krams, ASTIN: a Bayesian adaptive dose–response trial in acute stroke, *Clin. Trials J. Soc. Clin. Trials.* 2 (2005) 340–351, <https://doi.org/10.1191/1740774505cn094oa>.
- [17] M. Krams, K.R. Lees, W. Hacke, A.P. Grieve, J.-M. Orgogozo, G.A. Ford, Acute stroke therapy by inhibition of Neutrophils (ASTIN), *Stroke* 34 (2003) 2543–2548, <https://doi.org/10.1161/01.STR.0000092527.33910.89>.
- [18] A. Fisher, A. Andreasson, A. Chrysos, J. Lally, C. Mamasoula, C. Exley, J. Wilkinson, J. Qian, G. Watson, O. Lewington, T. Chadwick, E. McColl, M. Pearce, K. Mann, N. McMeekin, L. Vale, S. Tsui, N. Yonan, A. Simon, N. Marczin, J. Mascaro, J. Dark, An observational study of donor Ex Vivo lung perfusion in UK lung transplantation: DEVELOP-UK, *Health Technol. Assess.* 20 (2016) 1–276, <https://doi.org/10.3310/hta20850>.
- [19] J.A. Umans, P.H. Kloeg, J. Bronzwaer, The CAPTURE trial, *Lancet* 350 (1997) 445, [https://doi.org/10.1016/S0140-6736\(05\)64172-4](https://doi.org/10.1016/S0140-6736(05)64172-4).
- [20] Randomised placebo-controlled trial of abiciximab before and during coronary intervention in refractory unstable angina: the CAPTURE study, *Lancet* 349 (1997) 1429–1435, [https://doi.org/10.1016/S0140-6736\(96\)10452-9](https://doi.org/10.1016/S0140-6736(96)10452-9).
- [21] P.G. Casali, A. Le Cesne, A.P. Velasco, D. Kotasek, P. Rutkowski, P. Hohenberger, E. Fumagalli, I.R. Judson, A. Italiano, Y.C. Yang, J. Cuzick, S.M. Garland, W. Huh, S.K. Kjaer, O.M. Bautista, I.S.F. Chan, J. Chen, R. Gesser, E. Moeller, M. Ritter, S. Vuocolo, A. Luxembourg, A 9-valent HPV vaccine against infection and intraepithelial Neoplasia in women, *Obstet. Gynecol. Surv.* (2015), <https://doi.org/10.1097/OGX.0000000000000206>.
- [22] E.A. Joura, A.R. Giuliano, O.E. Iversen, C. Bouchard, C. Mao, J. Mehlsen, E.D. Moreira, Y. Ngan, L.K. Petersen, E. Lazcano-Ponce, P. Pitisuttithum, J.A. Restrepo, G. Stuart, L. Woelber, Y.C. Yang, J. Cuzick, S.M. Garland, W. Huh, S.K. Kjaer, O.M. Bautista, I.S.F. Chan, J. Chen, R. Gesser, E. Moeller, M. Ritter, S. Vuocolo, A. Luxembourg, A 9-valent HPV vaccine against infection and intraepithelial Neoplasia in women, *Obstet. Gynecol. Surv.* (2015), <https://doi.org/10.1097/OGX.0000000000000206>.
- [23] R. Kaplan, T. Maughan, A. Crook, D. Fisher, R. Wilson, L. Brown, M. Parmar, Evaluating many treatments and biomarkers in oncology: a new design, *J. Clin. Oncol.* (2013), <https://doi.org/10.1200/JCO.2013.50.7905>.
- [24] A.J. Atkinson, W.A. Colburn, V.G. DeGruttola, D.L. DeMets, G.J. Downing, D.F. Hoth, J.A. Oates, C.C. Peck, R.T. Schooley, B.A. Spilker, J. Woodcock, S.L. Zeger, Biomarkers and surrogate endpoints: preferred definitions and conceptual framework, *Clin. Pharmacol. Ther.* (2001), <https://doi.org/10.1067/mcp.2001.113989>.
- [25] SHEP Cooperative Research Group, Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension. Final results of the Systolic Hypertension in the Elderly Program (SHEP), *J. Am. Med. Assoc.* (1991), <https://doi.org/10.1001/jama.1991.03460240051027>.
- [26] B. Freidlin, E.L. Korn, Biomarker enrichment strategies: matching trial design to biomarker credentials, *Nat. Publ. Gr.* 11 (2013) 81–90, <https://doi.org/10.1038/nrclinonc.2013.218>.
- [27] C.N.A.M. Oldenhuis, S.F. Oosting, J.A. Gietema, E.G.E. de Vries, Prognostic versus predictive value of biomarkers in oncology, *Eur. J. Cancer* (2008), <https://doi.org/10.1016/j.ejca.2008.03.006>.
- [28] J.V. Carter, H.L. Roberts, J. Pan, J.D. Rice, J.F. Burton, N.J. Galbraith, M.R. Eichenberger, J. Jorden, P. Deveaux, R. Farmer, A. Williford, Z. Kanaan, S.N. Rai, S. Galandiuk, A highly predictive model for diagnosis of colorectal neoplasms using plasma MicroRNA, *Ann. Surg.* 264 (2016) 575–584, <https://doi.org/10.1097/SLA.0000000000001873>.
- [29] J. Fortea, M. Carmona-Iragui, B. Benejam, S. Fernández, L. Videla, I. Barroeta, D. Alcolea, J. Pegueroles, L. Muñoz, O. Belbin, M.J. de Leon, A.M. Maceski, C. Hirtz, J. Clarimón, S. Videla, C. Delaby, S. Lehmann, R. Blesa, A. Lleó, Plasma and CSF biomarkers for the diagnosis of Alzheimer's disease in adults with Down syndrome: a cross-sectional study, *Lancet Neurol.* 17 (2018) 860–869, [https://doi.org/10.1016/S1474-4422\(18\)30285-0](https://doi.org/10.1016/S1474-4422(18)30285-0).
- [30] M. Goshko, K. Nagashima, Y. Sato, Study Designs and Statistical Analyses for Biomarker Research, *Sensors* (Switzerland), (2012), <https://doi.org/10.3390/s120708966>.
- [31] H. Hao, D. Xiao, J. Pan, J. Qu, M. Egger, S. Waigel, M.G. Ann Sanders, W. Zacharias, S.N. Rai, K.M. McMasters, K.M. McMasters, A. Surg, Sentinel lymph node genes to predict prognosis in node-positive melanoma patients, *Oncol.* 24 (2017) 108–116, <https://doi.org/10.1245/s10434-016-5575-7>.
- [32] R. Simon, Advances in clinical trial designs for predictive biomarker discovery and validation, *Curr. Breast Cancer Rep.* 1 (2009) 216–221, <https://doi.org/10.1007/s12609-009-0030-4>.
- [33] J.C. Davies, C.E. Wainwright, G.J. Canny, M.A. Chilvers, M.S. Howenstine, A. Munck, J.G. Mainz, S. Rodriguez, H. Li, K. Yen, C.L. Ordoñez, R. Ahrens, Efficacy and safety of ivacaftor in patients aged 6 to 11 years with cystic fibrosis with a G551D mutation, *Am. J. Respir. Crit. Care Med.* (2013), <https://doi.org/10.1164/rccm.201301-0153OC>.
- [34] D.F. Hayes, R.C. Bast, C.E. Desch, H. Fritsche, N.E. Kemeny, J.M. Jessup, G.Y. Locker, J.S. Macdonald, R.G. Mennel, L. Norton, P. Ravdin, S. Taube, R.J. Winn, Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers, *JNCI J. Natl. Cancer Inst.* 88 (1996) 1456–1466, <https://doi.org/10.1093/jnci/88.20.1456>.
- [35] M. Antoniou, A.L. Jorgensen, R. Kolamunnage-Dona, Biomarker-guided adaptive trial designs in phase II and phase III: a methodological review, *PLoS One* 11 (2016) e0149803, <https://doi.org/10.1371/journal.pone.0149803>.
- [36] M. Antoniou, R. Kolamunnage-Dona, A.L. Jorgensen, Biomarker-guided non-adaptive trial designs in phase II and phase III: a methodological review, *J. Personalized Med.* 7 (2017), <https://doi.org/10.3390/jpm7010001>.
- [37] Biomarker-guided Trial Designs (BIGTeD) - an Online Tool to Help Develop Personalised Medicine, <http://www.bigted.org/>, Accessed date: 27 September 2018.
- [38] P. Tajik, A.H. Zwiderman, B.W. Mol, P.M. Bossuyt, Trial designs for personalizing cancer care: a systematic review and classification, *Clin. Cancer Res.* (2013), <https://doi.org/10.1158/1078-0432.CCR-12-3722>.
- [39] R. Simon, Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology, *Per. Med. Off.* 7 (2010) 33–47, <https://doi.org/10.2217/pme.09.49>.
- [40] R. Simon, Biomarker based clinical trial design, *Chin. Clin. Oncol.* (2014), <https://doi.org/10.3978/j.issn.2304-3865.2014.02.03>.
- [41] H.I. Scher, S.F. Nasso, E.H. Rubin, R. Simon, Adaptive clinical trial designs for simultaneous testing of matched diagnostics and therapeutics, *Clin. Cancer Res.* (2011), <https://doi.org/10.1158/1078-0432.CCR-11-1105>.
- [42] B. Dreno, J.F. Thompson, B.M. Smithers, M. Santinami, T. Jouary, R. Gutzmer, E. Levchenko, P. Rutkowski, J.-J. Grob, S. Korovin, K. Drucis, F. Grange, L. Machet, P. Hersey, I. Krajsova, A. Testori, R. Conry, B. Guillot, W.H.J. Kruit, L. Demidov, J.A. Thompson, I. Bondarenko, J. Jaroszek, S. Puig, G. Cinat, A. Hauschild, J.J. Goeman, H.C. van Houwelingen, F. Ulloa-Montoya, A. Callegaro, B. Dizier, B. Spiessens, M. Debois, V.G. Brichard, J. Louahed, P. Therasse, C. Debruyne, J.M. Kirkwood, MAGE-A3 immunotherapeutic as adjuvant therapy for patients with resected, MAGE-A3-positive, stage III melanoma (DERMA): a double-blind, randomized, placebo-controlled, phase 3 trial, *Lancet Oncol.* 19 (2018) 916–929, [https://doi.org/10.1016/S1470-2045\(18\)30254-7](https://doi.org/10.1016/S1470-2045(18)30254-7).
- [43] W. Jiang, B. Freidlin, R. Simon, Biomarker-adaptive threshold design: a procedure for evaluating treatment with possible biomarker-defined subset effect, *J. Natl. Cancer Inst.* (2007), <https://doi.org/10.1093/jnci/djm022>.
- [44] S. Matsui, Genomic biomarkers for personalized medicine: development and validation in clinical studies, *Comput Math Methods Med* (2013), <https://doi.org/10.1155/2013/865980>.
- [45] B. Freidlin, W. Jiang, R. Simon, The cross-validated adaptive signature design, *Clin. Cancer Res.* (2010), <https://doi.org/10.1158/1078-0432.CCR-09-1357>.
- [46] G.J. Kelloff, C.C. Sigman, Cancer biomarkers: selecting the right drug for the right patient, *Nat. Rev. Drug Discov.* (2012), <https://doi.org/10.1038/nrd3651>.
- [47] S.G. Baker, B.S. Kramer, D.J. Sargent, M. Bonetti, Biomarkers, subgroup evaluation, and clinical trial design, *Discov. Med.* 13 (2012) 187–192 <http://www.ncbi.nlm.nih.gov/pubmed/22463794>, Accessed date: 21 September 2018Accessed.
- [48] D.P. Byar, D.K. Corle, Selecting optimal treatment in clinical trials using covariate information, *J. Chronic Dis.* 30 (1977) 445–459, [https://doi.org/10.1016/0021-9681\(77\)90037-6](https://doi.org/10.1016/0021-9681(77)90037-6).
- [49] M.L. Rothenberg, D.P. Carbone, D.H. Johnson, Improving the evaluation of new cancer treatments: challenges and opportunities, *Nat. Rev. Canc.* 3 (2003) 303–309, <https://doi.org/10.1038/nrc1047>.
- [50] R. Simon, Clinical trials for predictive medicine, *Stat. Med.* 31 (2012) 3031–3040, <https://doi.org/10.1002/sim.5401>.
- [51] W.J. Lin, J.J. Chen, Class-imbalanced classifiers for high-dimensional data, *Briefings Bioinf.* (2013), <https://doi.org/10.1093/bib/bbs006>.
- [52] A. Callegaro, Futility for subgroup analyses in the adaptive signature design, *Stat. Methods Med. Res.* (2017), <https://doi.org/10.1177/0962280217739659>.
- [53] Z. Zhang, M. Li, M. Lin, G. Soon, T. Greene, C. Shen, Subgroup selection in adaptive signature designs of confirmatory clinical trials, *J. R. Stat. Soc. Ser. C Appl. Stat.* 66 (2017) 345–361, <https://doi.org/10.1111/rssc.12175>.
- [54] N. Stallard, T. Hamborg, N. Parsons, T. Friede, Adaptive designs for confirmatory clinical trials with subgroup selection, *J. Biopharm. Stat.* (2014), <https://doi.org/10.1080/10543406.2013.857238>.
- [55] J.J. Chen, S.-J. Wang, C.-A. Tsai, C.-J. Lin, Selection of differentially expressed genes in microarray data analysis, *Pharmacogenomics J.* 7 (2007) 212–220, <https://doi.org/10.1038/sj.tpj.6500412>.
- [56] W.S. Noble, How does multiple testing correction work? *Nat. Biotechnol.* (2009), <https://doi.org/10.1038/nbt1209-1135>.

- [57] T.-T. Vo, A. Vivot, R. Porcher, Impact of biomarker-based design strategies on the risk of false-positive findings in targeted therapy evaluation, *Clin. Cancer Res.* (2018), <https://doi.org/10.1158/1078-0432.CCR-18-0328> clincanres.0328.2018.
- [58] A.C. Cambon, K.B. Baumgartner, G.N. Brock, N.G.F. Cooper, D. Wu, S.N. Rai, Classification of clinical outcomes using high-throughput informatics: Part 2 - parametric method reviews, *Model Assist. Stat. Appl.* (2015), <https://doi.org/10.3233/MAS-140321>.
- [59] A.C.C. Cambon, K.B.B. Baumgartner, G.N.N. Brock, N.G.F.F. Cooper, D. Wu, S.N.S.N. Rai, Classification of clinical outcomes using high-throughput informatics: Part 2 - parametric method reviews, *Model Assist. Stat. Appl.* (2015), <https://doi.org/10.3233/MAS-140321>.
- [60] A.C. Cambon, K.B. Baumgartner, G.N. Brock, N.G.F.F. Cooper, D. Wu, S.N. Rai, Properties of adaptive clinical trial signature design in the presence of gene and gene-treatment interaction, *Commun. Stat. Simulat. Comput.* 46 (2017) 8233–8250, <https://doi.org/10.1080/03610918.2016.1275690>.
- [61] H. Tian, K. Liu, Biomarker enrichment considerations in oncology early development single-arm studies, *J. Biopharm. Stat.* 28 (2018) 282–291, <https://doi.org/10.1080/10543406.2017.1379533>.
- [62] R. Simon, S.J. Wang, Use of genomic signatures in therapeutics development in oncology and other diseases, *Pharmacogenomics J.* (2006), <https://doi.org/10.1038/sj.tpj.6500349>.
- [63] Y. Choai, S. Matsui, Estimation of treatment effects in all-comers randomized clinical trials with a predictive marker, *Biometrics* 71 (2015) 25–32, <https://doi.org/10.1111/biom.12253>.
- [64] E.L. Korn, B. Freidlin, Reader reaction on estimation of treatment effects in all-comers randomized clinical trials with a predictive marker, *Biometrics* 73 (2017) 706–708, <https://doi.org/10.1111/biom.12592>.
- [65] G. Mi, Enhancement of the adaptive signature design for learning and confirming in a single pivotal trial, *Pharmaceut. Stat.* 16 (2017) 312–321, <https://doi.org/10.1002/pst.1811>.
- [66] B. Freidlin, E.L. Korn, Biomarker-adaptive clinical trial designs, *Pharmacogenomics* (2010), <https://doi.org/10.2217/pgs.10.153>.