All-Comers versus Enrichment Design Strategy in Phase II Trials

Sumithra J. Mandrekar, PhD, and Daniel J. Sargent, PhD

Abstract: Designs for biomarker validation have been proposed and used in the phase III oncology clinical trial setting. Broadly, these designs follow either an enrichment (i.e., targeted) strategy or an all-comers (i.e., unselected) strategy. An enrichment design screens patients for the presence or absence of a marker or a panel of markers and then only includes patients who either have or do not have a certain marker characteristic or profile. In contrast, all patients meeting the eligibility criteria (regardless of a particular biomarker status) are entered into an all-comers design. The strength of the preliminary evidence, the prevalence of the marker, the reproducibility and validity of the assay, and the feasibility of real-time marker assessment play a major role in the choice of the design. In this report, we discuss the parameters under which the enrichment or an all-comers design strategy would be appropriate for phase II trials.

Key Words: All-comers, Adaptive, Biomarker, Enrichment, Phase II, Randomized.

(*J Thorac Oncol.* 2011;6: 658–660)

Phase I clinical trials are first-in-man studies that primarily focus on the safety profile of a regimen. With the development of molecularly targeted therapies, vaccines, and immunotherapy, the focus of phase I trials is shifting toward the identification of a biologically optimal dose, as these agents typically have limited toxicity concerns but unknown dose-efficacy relationships.[1] Most phase I trials of targeted agents include an expanded cohort of patients with a certain marker profile at the recommended dose to understand the treatment effects (efficacy or toxicity) within marker defined subgroups. An example is the testing of crizotinib in patients with anaplastic lymphoma kinase (ALK) -positive non-small cell lung cancer (NSCLC).[2,3] An expanded cohort of patients with ALK-positive tumors (82/1500 patients, ~5% of patients who underwent screening) was enrolled into the phase I trial. The observed response rate was 57% per the RECIST, and disease control rate at 8 weeks was 87%. Based on these data, this agent is currently being tested in a phase III setting for patients with ALK-positive NSCLC. Such exceptional results are, however, rare, and most agree to proceed to phase II testing after completion of phase I.

Phase II clinical trials are designed primarily to identify promising experimental therapies that can then be tested further in a definitive phase III trial. Although there is a wealth of literature on the types of phase II designs (single arm, randomized, screening, selection, etc.)[4] and under what circumstances a particular design is valid, there is very little published on the question of patient selection (i.e., biologic subsetting) when evaluating molecularly targeted agents.

In this brief report, we first discuss the role of randomization and adaptive design strategies in the context of assessing targeted therapies. Next, we review the concept of enrichment (or targeted) and all-comers (or unselected) designs in the setting of phase II trials and discuss the relevance of each in light of the strength of the preliminary evidence, the prevalence of the marker, reproducibility and validity of the assay, and the feasibility of real-time marker assessment.

RANDOMIZATION AND ADAPTIVE DESIGN STRATEGIES

Randomized phase II trials are critical in the current era of constrained resources and availability of multiple promising therapies for a given disease. Changes in patient populations based on biologic subsetting, evolution in imaging technologies, make comparison against historical controls inaccurate. Moreover, randomized phase II designs are essential for biomarker discovery, specifically for making the distinction between a prognostic and predictive marker.[4,5] Adaptive design strategies are a class of randomized phase II designs by which a variety of marker signatures and drugs can be tested. In these designs, the success of the drug-biomarker subgroup is assessed on an ongoing manner, and in some designs, the randomization ratio is altered and/or the underperforming drugs and/or the biomarker subgroups tested eliminated midway through the trial. The essential requirements for these designs are (1) a rapid and reliable endpoint, which is somewhat challenging as most oncology trials use time to event endpoints or endpoints that involve following a patient's status for a predetermined time period (such as the progression status at 2 years) and (2) real-time access to data, which can be a daunting task in multicenter trials.

Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota.

Disclosure: The authors declare no conflicts of interest.

Address for correspondence: Sumithra J. Mandrekar, PhD, Department of Health Sciences Research, Mayo Clinic, 200 1st ST SW, Rochester, MN 55905. E-mail: mandrekar.sumithra@mayo.edu

ISSN: 1556-0864/11/0604-0658

TABLE 1. Parameters to Consider for Choice of a Design

Criteria	Design	
	Enrichment	All Comers
Preliminary evidence		
Strongly suggest benefit in marker-defined subgroups	Yes	—
Uncertain about benefit in overall population vs. marker-defined subgroups	—	Yes
Assay reproducibility and validity		
Excellent (high concordance between local and central testing, commercially available kits, etc.)	Yes	—
Questionable	—	Yes
Turnaround times		
Rapid (2–3 days; without causing delay in the start of therapy)	Yes	Yes
Slow to modest (1 week or more)	—	Yes (retrospective marker subgroup assessment)

Two phase II trials that used an adaptive design strategy are I-SPY 2 (investigation of serial studies to predict therapeutic response with imaging and molecular analysis 2) and BATTLE (biomarker-integrated approaches of targeted therapy for lung cancer elimination trial).[6–8] I-SPY 2 is an ongoing neoadjuvant trial in breast cancer that is designed to compare the efficacy of standard therapy to the efficacy of novel drugs in combination with chemotherapy. All drugs will be evaluated against biomarker signatures; regimens that have a high predictive probability of being effective will be moved forward to phase III testing with their corresponding biomarker signature(s). Regimens that have a low probability of efficacy for all biomarker signatures will be dropped from further development.[6] BATTLE used an outcome-based adaptive randomization design for randomizing patients to treatment choices based on multiple biomarker profiles in NSCLC. Patients had their tumors tested for 11 different biomarkers and were subsequently randomized to one of four treatment choices. The first 97 patients enrolled to the trial were randomly assigned to one of the four treatments equally, and subsequently, new patients were adaptively randomized. The preliminary results from BATLLE are promising, wherein, as hypothesized, each drug was found to work best for patients with a specific molecular profile.[7,8] Two successor trials, BATTLE 2 and BATTLE 3, are currently in development.

ENRICHMENT VERSUS ALL-COMERS DESIGNS

An enrichment design screens patients for the presence or absence of a marker or a panel of markers and then only includes patients in the clinical trial who either have or do not have a certain marker characteristic or profile.[9] This results in a stratification of the study population, with a goal of understanding the safety, tolerability, and clinical benefit of a treatment in the subgroup of the patient population defined by a specific marker status. This design is based on the paradigm that not all patients will benefit from the study treatment under consideration but rather that the benefit will be restricted to a subgroup of patients who either express or do not express a specific molecular feature. In an all-comers design, all patients meeting the eligibility criteria, which does not include a specific status on the biomarker in question, are entered.[5] The ability to provide adequate tissue may be an eligibility criterion for these designs but not the specific biomarker result or the status of a biomarker characteristic. The need to collect upfront tissue and blood specimens is critical in the current era of targeted therapies to match the right patient to the right drug.[10]

Table 1 lists some general criteria to consider when deciding between enrichment versus all-comers design in a phase II setting. Enrichment designs are clearly appropriate when the marker prevalence is low ($<$15–20%), whereby it is not feasible to include all patients regardless of the marker status as the treatment effect in the overall population will be diluted. In this case, an all-comers trial design would require a prohibitively large sample size. When the marker prevalence is high ($\geq$50%), the assay performance is not well established (no established cutpoint for marker status definition, laboratory is not Clinical Laboratory Improvement Amendments certified, etc.), the turn around times for marker assessment are long (more than a week, for example, in second- or third-line treatment settings), and the preliminary evidence is unclear, an all-comers design is appropriate. In most instances, however, an all-comers design should clearly incorporate a prospectively specified subgroup analyses for retrospective assessment of the treatment effect within biomarker-defined subgroups. This is critical to ensure that the effect of the drug is tested both on the overall and prospectively defined subsets of patients so as to not incorrectly conclude that the drug is not effective, when it may be effective for a smaller subset of the population.

In cases where the prevalence of the marker in question is moderate (between 20% and 50%), then a possible strategy could be as follows:

1. Perform a single-arm enrichment trial (pilot) as a proof of concept that the treatment likely has a major effect within a marker subgroup.
2. Perform an unselected phase II (randomized) trial, either.
   a. Stratified by marker status, with a primary hypothesis within the marker subgroup hypothesized to derive the most benefit based on the pilot trial in step 1, but

sufficient patients to potentially demonstrate lack of benefit in the other subgroup(s).

b. Adaptive design by evaluating success in an ongoing manner within the marker subgroups.

In summary, enrichment designs are appropriate when one or more of the following are likely:

1. The treatment in question has modest absolute benefit in the unselected population but can cause significant toxicity.
2. In the absence of selection, therapeutic results are similar whereby an enrichment design (even if incorrect) would not hurt.
3. There is compelling preliminary evidence to suggest that patients with or without that marker profile do not benefit from the treatments in question, thus including all patients regardless of the marker profile is ethically not possible.
4. Assay reproducibility and accuracy is well established, for example, a high degree of concordance between local and central laboratory testing.
5. Rapid turnaround time for marker assessment is available, so as to not delay treatment initiation.

SUMMARY

Stewart et al.[11] demonstrated through well-conducted simulation studies the impact of subpopulation characteristics on overall study outcomes. One of their main conclusions is that although molecular profiling is expensive, not doing so is ultimately far more expensive and gives the wrong answer. Clearly, biomarker identification is critical to future ongoing oncology drug development. In the setting of phase II trials, when the mechanism of action of a targeted agent is known (i.e., the target is known, and it is well established that the agent under investigation selectively inhibits or activates the target), then clinical trials with an enrichment strategy are appropriate, such that large gains can be expected in a subset of patients. On the other hand, an all-comers design is optimal where preliminary evidence regarding the treatment benefit and/or the assay performance characteristics is uncertain, to allow a more comprehensive assessment of a new agent's potential activity.

ACKNOWLEDGMENTS

Supported, in part, by the National Cancer Institute Grants: Mayo Clinic Cancer Center (CA-15083) and the North Central Cancer Treatment Group (CA-25224).

REFERENCES

1. Mandrekar SJ, Qin R, Sargent DJ. Model-based phase I designs incorporating toxicity and efficacy for single and dual agent drug combinations: methods and challenges. *Stat Med* 2010;29:1077–1083.
2. Hallberg B, Palmer RH. Crizotinib—latest champion in the cancer wars? *N Engl J Med* 2010;363:1760–1762.
3. Kwak EL, Bang Y-J, Camidge DR, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 2010;363:1693–1703.
4. Mandrekar SJ, Sargent DJ. Randomized phase II trials: time for a new era in clinical trial design. *J Thorac Oncol* 2010;5:932–934.
5. Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. *J Clin Oncol* 2009;27:4027–4034.
6. Barker AD, Sigman CC, Kelloff GJ, et al. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin Pharmacol Ther* 2009;86:97–100.
7. Zhou X, Liu S, Kim ES. Bayesian adaptive design for targeted therapy development in lung cancer-a step towards personalized medicine. *Clin Trials* 2008;5:181–193.
8. Kim ES, Herbst RS, Lee JJ, et al. The BATTLE trial (Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination): personalizing therapy for lung cancer [abstract]: 101 Annual Conference of American Association for Cancer Research, Washington, DC, 2010, abstract LB-1.
9. Simon R, Maitournam A. Evaluating the efficiency of targeted designs for randomized clinical trials. *Clin Cancer Res* 2004;10:6759–6763.
10. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst* 2009;101:1–7.
11. Stewart DJ, Whitney SN, Kurzrock R. Equipoise lost: ethics, costs, and the regulation of cancer clinical research. *J Clin Oncol* 2010;28:2925–2935.