The landscape of non–small cell lung cancer (NSCLC) diagnosis and treatment has dramatically changed in the past few years owing to the successful pairing of biomarker-defined cohorts of patients with targeted therapeutics: namely, *EGFR* mutations as biomarkers of benefit from epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKI) and *EML-ALK* translocations as biomarkers of benefit from ALK TKIs (1, 2). Testing NSCLC patients for these and other biomarkers at the time of diagnosis is becoming more routine because it affects decisions about treatment as well as patient outcomes (3). These examples also underscore the value of exploring biomarkers during early clinical trials with targeted therapies. However, most novel targeted therapies studied in NSCLC clinical trials are not administered as initial therapies but rather as second, third, or later lines of treatment. Clinical trial designs frequently do not mandate tumor tissue from all patients but attempt post hoc analyses of biomarker status among those with available tissue from the original diagnostic biopsy, typically 25% to 45% of the study population (1, 4, 5). Not only is this “strategy of convenience” not comprehensive, but it also risks inaccurate conclusions if intervening treatments have altered the biologic and/or biomarker status since the time of the archival biopsy specimen. Capturing the biomarker status for all participants at the time of drug administration maximizes the chances of discovering the relationship between putative biomarkers and response to novel treatments. In this issue of Cancer Discovery, Kim and colleagues (6) describe their landmark effort to accomplish these goals within the framework of the Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE) trial.

BATTLE was designed as an umbrella structure, within which 4 separate therapeutic clinical trials for NSCLC were nested: erlotinib, vandetanib, erlotinib plus bexarotene, and sorafenib. All NSCLC patients entering the BATTLE program (at the time of second-line therapy or beyond) first underwent a core tumor biopsy for the purpose of obtaining up-to-date biomarker status and then were randomized to one of the 4 treatment arms. Biomarker analysis was done in real time, with a large panel of mutation, gene copy number, and immunohistochemistry analyses performed on each sample. The results of these studies grouped patients into predefined biomarker signature groups, which could then be used to evaluate treatment–biomarker interactions. The initial 97 patients were randomized equally into the 4 different treatments, with a 23% chance of being placed into each treatment arm. Results from these 97 patients were then assessed, comparing the outcome of interest—8-week disease control rate (DCR)—with the biomarker status within each treatment arm. With this information, future randomization probabilities were adjusted (rather than being equal) using a Bayesian model. This adjustment means that if a patient was found to have a particular biomarker signature on biopsy, he or she would have a >25% chance of being randomized to a treatment on which prior patients with the same biomarker signature had done well with respect to 8-week DCR. The pattern continually repeated, so that the more patients with a particular signature did well on a particular therapy, the higher the probability of being assigned to that therapy for subsequent similar patients. The authors planned to benchmark the BATTLE 8-week DCR against the historical rate of 30% for similar patients. The study was not designed to determine if significant associations existed between particular biomarkers and treatments.

The innovative adaptive randomization design attempts to increase the opportunity for each patient to receive the most effective experimental treatment possible, a feature that is attractive to potential patients and their oncologists alike. However, this type of Bayesian design has not yet been used in clinical research with the frequency required to establish robust standard practices. Some concerns have arisen that adaptive randomization may worsen the precision of estimates of treatment effect by increasing variability, mainly owing to unequal subject allocation to the treatment arms (7). The potential caveats of the Bayesian design aside, the authors deserve tremendous praise for accomplishing this Herculean task. The difficulty in establishing the infrastructure and multidisciplinary collaborations necessary to successfully carry out 255 core needle biopsies with real-time multiplexed genotype and other biomarker analyses cannot be overstated.
This study showed that these procedures were safe (<1% incidence of serious complications among patients undergoing lung biopsy) and that real-time biomarker assessment is possible (83% of patients could be categorized into one of the predefined biomarker signature cohorts). Had the authors designated feasibility as an end point for the BATTLE trial, they would likely have met their benchmarks. The University of Texas MD Anderson Cancer Center (Houston, Texas) has shown that trials mandating pretreatment biopsies coupled with complex real-time biomarker analysis are feasible.

The primary end point of the trial, 8-week DCR, has been shown previously to be a reasonable surrogate for overall survival (8). The authors chose this somewhat unconventional end point because the adaptive randomization design requires an end point that can be rapidly determined for each patient, to facilitate the Bayesian algorithm going forward. The overall 8-week DCR was 46% among 244 evaluable patients. When examining 8-week DCR by treatment arm, they observed 34% for erlotinib, 33% for vandetanib, 50% for erlotinib + bexarotene, and 58% for sorafenib. The study did not report the 8-week DCR among the initial cohort of patients assigned to treatment by equal randomization because, as the authors point out, the study was not designed or powered to assess this crucial question. However, these data would be useful to determine whether the Bayesian trial design truly affected patient outcomes and steered patients toward the most effective therapies. Because these initial observations formed the basis for adaptive randomization used in the rest of the trial, an appreciation of the magnitude of variation among the initial cohort would increase general confidence in the potential benefits of this adaptive randomization study design.

The BATTLE investigators did report treatment efficacy by biomarker signature groups. The results from the large number of biomarkers assessed on each patient were condensed into 4 biologically relevant groups, and the biomarker-positive versus biomarker-negative couples were then analyzed by treatment arm. Significant biomarker–treatment relationships were defined as those in which the biomarker-positive group had an 80% probability of achieving better outcomes than the historical 8-week DCR of 30%. This threshold turned out to have a relatively low sensitivity, as 8 of the 20 pairs they examined had a significant biomarker–treatment relationship. However, setting the bar low is acceptable in this sort of hypothesis-generating exercise and allows the investigators to observe correlations that might not have been envisioned a priori. In fact, one of the most interesting findings from the BATTLE study is the promising biomarker–treatment relationship between KRAS or BRAF mutation–positive patients and sorafenib therapy (7% of the biomarker-positive patients had 8-week DCR). It seems likely that the adaptive randomization design enhanced the ability to make this observation by aggregating these patients into the sorafenib arm. A total of 27 patients were positive for the KRAS/BRAF biomarker signature, and half of them (n = 14) were assigned to sorafenib. Examined from the opposite perspective, the BATTLE design directed KRAS-positive patients away from the other 3 treatment arms, which all included EGFR inhibitors; this strategy was likely beneficial because previous data have demonstrated that EGFR TKIs are ineffective against KRAS mutant cancers (5, 9, 10). Although BATTLE does not definitively confirm a relationship between KRAS mutations and sorafenib efficacy, it certainly provides the impetus for ongoing validation studies as well as future studies to assess more potent inhibitors of the RAF/MEK/ERK signaling axis in these cancers.

It is notable that decisions about how to define biomarker groupings at the outset of the trial may have obscured the ability to make some observations. As the authors point out, the most obvious example of this is the pairing of EGFR mutations with EGFR gene copy number into a single group. At the time BATTLE was designed, this seemed a reasonable strategy because both biomarkers were considered predictive for benefit from EGFR TKIs. However, it has now become clear that EGFR mutations are much more reliable for identifying patients who will benefit from EGFR TKI therapy (11), and the aggregation with EGFR gene copy number likely muted the relationship observed between the EGFR biomarker group and erlotinib-based therapies. It is not known at this time if other grouping decisions may have influenced the observations in a similar way. This point highlights the concern that a large and complex research structure such as BATTLE, with an umbrella framework and multiple nested treatment studies, does not have the agility to adapt quickly as knowledge outside the trial advances. For example, the ALK translocation story developed all the way from the bench to the bedside during the course of the BATTLE trial (2, 12). The investigators acknowledge this development and plan to exclude patients with ALK translocations from future BATTLE studies so that they will be steered toward ALK-directed therapies; one hopes that other such biomarker–treatment combination success stories will develop over the course of subsequent BATTLE studies, and the investigators will need to be mindful to build strategies for identifying and triaging these patients into their future designs.

It remains challenging to define the exact circumstances in which the Bayesian trial design will more rapidly identify clinically significant biomarker–targeted therapy relationships, compared with the more common clinical trial designs, such as prospective biomarker-directed trials (e.g., crizotinib in ALK-translocated cancers) or careful retrospective examination of specific biomarkers in conventional targeted therapy trials (13–15). The BATTLE study design may provide a more distinct advantage in the study of novel drugs without clearly understood mechanisms of action or in situations when the biologic characteristics of the target are uncertain. In these situations, real-time, complex biomarker analyses may accelerate the identification and further testing of potential biomarker–therapy relationships. Conversely, when hypotheses about the drug target and its biologic features are well understood, the adaptive randomization strategy may be less efficient than either prospective biomarker-directed trials addressing mature hypotheses or conventional targeted therapy studies in less restricted patient populations that incorporate retrospective analyses of well-defined biomarkers.

In summary, the BATTLE trial offers proof that, with a concerted, determined effort, we can successfully raise the bar for lung cancer clinical trials research to include
comprehensive pretreatment biopsies and genotyping for all participants. We believe that such efforts have great potential to exponentially increase our understanding of patients who benefit from targeted therapies and are likely to accelerate and improve the drug development process. Although it remains to be determined when an adaptive randomization design such as the one used by BATTLE investigators ultimately increases the efficiency of discovery, these investigators have clearly set a new standard for acquiring tissue and performing comprehensive biomarker evaluation in real time. The rest of us will have no valid excuses when future compelling trials demand the same.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Published OnlineFirst April 3, 2011.

REFERENCES
