RTI (h)(s)
Health Solutions

WHITE PAPER

Adaptive Clinical Trial Design

James Rosenzweig and David McSorley

Clinical trials in the pharmaceutical industry are incorporating adaptive design methods with greater frequency, as the economic resources needed for clinical research have expanded. These designs leverage accumulating information during a trial in real time and have the potential to reduce the costs and streamline the time frames for clinical trials in drug development, particularly in the earlier phases during proof of concept and dose selection.

Adaptive design is usually defined as the use of accumulating data obtained during the conduct of the trial to decide how to modify aspects of the trial as it progresses, without affecting its validity and integrity (Gallo et al., 2006). This form of research design involves a prospectively planned opportunity to modify the parameters of the study based on analysis of interim data while the study is being performed. The prospective planning of key time points for interim analysis, and the parameters for change in the study, have to be set before the study is underway.

Adaptive clinical trial designs were further characterized by the US Food and Drug Administration (FDA) (FDA, 2010). The FDA divides adaptive designs into two categories—"well understood" and "less well understood." Those considered well understood have a record of being performed in the past, with established statistical methods and familiarity with the FDA from previously approved studies. Less well understood designs fail to meet these criteria.

According to the FDA guidelines, an adaptive clinical trial can involve:

- Analysis at decision points during the trial to stop or to adjust patient accrual
- Interim evaluations to determine if the trial should be stopped early—because of a determination of success, demonstration of futility, or finding of unacceptable harm to subjects
- Hypothesis reversal of noninferiority to superiority or vice versa
- Discontinuation of arms or doses, or the changing of doses while the trial is underway
- Modification of the randomization rate to increase the probability that a patient is allocated to the most appropriate arm

Others have argued for a more liberal definition of adaptive design, which allows not only for prospective adaptations, but for concurrent (ad hoc) and retrospective adaptations (Chow, 2014). The use of Bayesian methodologies can enable greater insight on which options for design changes should be made during the course of the study. Adaptive designs use interim analyses of the accumulating data from within an ongoing study to modify various aspects of the trial and then continue under the modifications. The different types of adaptations in study design have often been put into categories, which helps clarify the specific issues to be dealt with (Chow and Chang, 2008). However, categories may overlap and studies may combine multiple strategies, as seen in the following:

1. Adaptive randomization

These designs allow for changing the randomization schedules of a study by adjusting the probability of treatment assignment based on prior assignments in order to either avoid an imbalance of important patient characteristics between treatment groups or to increase the likelihood of being assigned to a particular treatment group. Adaptive randomization schemes include treatment-adaptive (TA), covariate-adaptive (CA), and response-adaptive (RA) randomization. Treatment-adaptive

and covariate-adaptive designs aim to balance the treatment groups with respect to patient characteristics by changing the way the subsequent patients are assigned to a treatment group, whereas response-adaptive randomization aims to increase the probability of being assigned to the treatment group with more favorable responses. Response-adaptive schemes can therefore cause imbalances in patient characteristics that may require subsequent adjustment for these imbalances resulting in a combined RACA design (Ning and Huang, 2010). Adaptive randomization schemes have the most utility in small ($n < 100$), early-phase trials—where equal probability randomization may not produce the desired balance in patient characteristics among treatment groups—as the designs quickly become impractical for large or longer-duration trials.

2. Adaptive group sequential design

Classical group sequential methods use repeated significance testing on accumulating groups of enrolled patients to decide whether to stop or continue a trial based on established stopping boundaries for each test that maintain the overall type I and type II error rates across all tests. Type I error occurs when the null hypothesis is true and type II error occurs when the null hypothesis is false but is not rejected. The concept of adaptive design allows for additional changes to a study protocol as it progresses as a result of analysis of interim data (Bauer and Köhne, 1994). These include potential modification, deletion or addition of treatment arms, re-estimation of the sample size, change of study endpoints, changing of dose and/or duration of treatment, and modification of randomization schedules. Adaptive group sequential designs combine the concepts of both early stopping and re-engineering of the design based on the observed early results.

An example of a group sequential design that employed adaptive elements was the Diabetes Control and Complications Trial (DCCT) (DCCT, 1993) and its follow-up, the Epidemiology of Diabetes Interventions and Complications (EDIC) study (Nathan et al., 2005). The DCCT was a multicenter, randomized clinical trial designed to determine whether or not very tight control of blood glucose (experimental therapy at the time) could reduce or prevent the microvascular and macrovascular complications of type 1 diabetes, when compared with the standard insulin treatment of the day. The study had parallel arms consisting of subjects with no complications and those with very early complications. The study was initiated in 1982. A feasibility phase with 278 subjects was completed in 1985, and an interim analysis determined that the experimental intervention was safe and effective in improving glycemic control by a sufficient margin to allow for full recruitment, which expanded to 1,441 subjects. A specific concern had been that the intervention to intensify glycemic control would be accompanied by an unacceptably high risk of severe hypoglycemia; this was found not to be the case. An independent data, safety, and quality committee (DSQ) followed the results on a regular basis while the investigators were blinded.

The full randomized controlled clinical trial phase was stopped prematurely after a mean follow-up time of 6.5 years, when the benefits of intensive treatment were found to be incontrovertible by the DSQ and not likely to be reversed over time. At that point, subjects on intensive control were encouraged to continue and those originally assigned to conventional treatment were advised to switch to intensive treatment. During the closeout phase of the trial, they were provided the resources and education to implement this. A total of 1,375 patients agreed to participate in the EDIC study, the open-label follow-up to the DCCT, which continues to this day. The EDIC study continues to monitor the DCCT patients for level of complications and cardiovascular events. Although the level of glycemic control in the two groups came together not long after the randomized phase was completed, it was found that those in the intensive arm for only 6.5 years continued to have fewer microvascular complications for at least 30 years of follow-up, and macrovascular benefits as well.

3. Flexible sample size re-estimation

This design enables the size of the sample in the study to be changed or re-estimated based on unblinded interim effect size data and often may be included as one of the adaptations in an adaptive group sequential design. In a fixed-sample study design, the sample size is determined before the study and is based on prior estimates of the clinically meaningful effect size between the treatment and control groups that can be achieved for a specified power and type I error rate. It is not uncommon for effect sizes to be initially specified incorrectly, resulting in an underpowered design, especially if the variability turns out to be larger than initially specified. As a result, it may be desirable to adjust the sample sizes based on accrued data while a trial is underway. However, any sample size re-estimation should be planned in advance and done using appropriate group sequential methods so as to preserve the type I error rate.

4. "Drop the losers"

When multiple treatment arms are used, it is often helpful to have a multistage design to enable the investigators to drop arms that are shown to be inferior to others. This design is sometimes referred to as selection design or "pick up the winners,"

as it also allows adding additional arms (Chow, 2014). Typically, it is the first stage of a two-stage design, in which the inferior arms are dropped according to criteria specified in the beginning of the study. The winning treatment groups go on to the next stage of the study. It is also possible to use different analytic approaches (e.g., Bayesian predictive probabilities vs. frequentist hypothesis testing) for the progression criteria between stages, so the study needs to be designed to have sufficient power for those stages using a frequentist hypothesis testing approach (e.g., at the end of the trial). It is possible that dropping or adding the wrong dose groups could lead to loss of valuable information that would have been helpful at the end of the study. Because of this, it is important to use valid and well-considered decision rules or criteria for selecting dose groups.

5. Adaptive dose finding

Study designs that employ adaptive finding are often used in phase 1 or 2 studies in order to determine the maximum tolerated dose (MTD) of the medication, which is often used as the optimal dose for later-phase clinical studies. It helps to avoid having too many subjects exposed to dose-limiting toxicity (DLT), and a small number of subjects can be used to identify the MTD. To achieve this end, careful selection of the appropriate initial dose is important, as well as the dose range and parameters for dose escalation or dose reduction.

In early-phase oncology studies, it is often difficult to balance toxicity with clinical effectiveness. One dose-finding method commonly used the 3+3 design. With this procedure, three subjects receive a particular dose of study medication. If no patients experience DLT, then the dose is increased by a predetermined amount. If two or more experience DLT, the dose is decreased. If one subject experiences DLT, then three subjects are added. If only one of the six subjects experiences DLT, the dose can be increased, and if two or more have DLT, then the dose is decreased.

Increasingly, dose-finding studies have utilized an iterative model-fitting process, often called the continuous reassessment method (CRM), to find the MTD. A number of studies have shown though simulations that CRM model-based designs are more accurate and effective than the 3+3 design. They are able to determine the MTD more quickly, and a greater percentage of the subjects treated in these studies are found to be at or near the MTD. These newer CRMs employ dose-escalation algorithms that emphasize overdose control, or that optimize the time to event or late-dose toxicities to refit the dose-toxicity curve after each dose level's toxicity outcome is observed (Garrett-Mayer, 2006). In general, it has been shown that these designs do not pose major safety concerns. On average about 25 to 35 subjects are required to test 5 or 6 dosage levels (Iasonos and O'Quigley, 2014).

6. Adaptive treatment switching

In this situation, the design of the study can permit the investigator to switch the patient to an alternative treatment if there is evidence of lack of efficacy, progression of disease, or safety problems with the initial medications. This is commonly used in oncology trials because of compassion issues related to the consequences of withholding a possible beneficial treatment. The statistical analysis must also adjust for the treatment switching. In an evaluation of nine methods that adjust for treatment switching, Fox et al. (2011) found that only the rank-preserving structural failure time (RPSFT) model of Robins and Tsiatis (1991) and the iterative parameter estimation (IPE) method of Branson and Whitehead (2002) gave accurate and consistent results, with an advantage to the IPE method.

7. Adaptive hypothesis design

It is possible to make potential changes to the hypothesis of a study based on interim data that is collected. This can be done by applying the closure principle (Marcus et al., 1976) to the hypotheses of interest and testing each of them by using an appropriate combination test (Bretz et al., 2006). Some examples include changing from a single hypothesis to multiple hypotheses or a composite hypothesis, switching from a superiority hypothesis to a noninferiority hypothesis, switching between the null and alternative hypotheses, or changing the primary and secondary endpoints.

8. Phase 1-2 or phase 2-3 seamless trial design

Adaptive seamless design is used to combine the aims and objectives of what would normally be considered separate trials into one study. Most likely, a phase 1 study would be combined with a phase 2 study of the same compound. Similarly, phase 2 and phase 3 studies can be combined. Typically, a phase 1 trial to establish the MTD of a drug can be combined with an early phase 2a trial to investigate the efficacy of the drug at that dose. A phase 2b dose-ranging study can be combined with a confirmatory phase 3 trial with more subjects and investigational sites, and perhaps different endpoints. It could be set up as a two-stage study, with the interim analysis serving as a decision point for whether the trial should be stopped or expanded. With a seamless phase 2/3 design, valuable information can be obtained in the first stage that could help in decisions made during the conduct of the second stage, in particular which dose(s) to retain in stage 2. Because this design would allow for use of data acquired from both

stages, there can be some economies of scale. Costs can be saved through combining of evidence across the two stages. Sample size can be reduced in comparison to running two separate studies. There would be no lead time between the two studies so that time can be saved. Instead of starting anew with institutional review board approval, site recruitment, and subject enrollment, the process would be expanded seamlessly in the second phase of the study. However, extra planning is necessary and the statistical methodology must account for potential biases and multiple looks at the data, and how to combine the data from the different stages to make sure that the overall validity of the study can be maintained.

9. Biomarker-adaptive design

Biomarkers may be collected in some studies to detect normal function or assess pathogenic or pharmacological processes in response to the therapeutic agent under investigation. However, they should not be confused with primary endpoints. A biomarker that correlates well with a clinical endpoint can be considered a prognostic biomarker. These can be used to identify information about the natural course of the disease being studied, irrespective of whether the subject is randomized to the treatment in question. At the start of the study, prognostic biomarkers can be used to stratify patients by good or poor prognosis or disease severity, for efficiency of recruiting, or for subgroup analyses to identify the degree of expected responsiveness or sensitivity to the treatment being studied. They should not be used to select the particular treatment under study. Biomarkers can be used to identify a particular therapy for use during the study in affected patients. They are most often used in exploratory studies to identify the appropriate criteria for patients to be selected for participation in later trials. However, if it is imbedded in a trial to modify patient eligibility after the interim analysis, the statistical methodology must account for how the data collected before and after the interim analysis will be combined and analyzed at the end of the study.

An example of this design has been the "Biomarker-integrated Approaches to Targeted Therapy of Lung cancer Elimination (BATTLE) trial (Kim et al., 2011). This trial included patients with stage IV recurrent non–small-cell lung cancer. The primary endpoint was the 8-week disease control rate (DCR). Four biomarker profiles were examined, and four different drug therapies were employed, with one therapy targeting each biomarker profile. The trial looked at the four biomarkers and aimed to identify their predictive roles in providing better treatment efficacy in terms of the DCR to patients in the trial based on their biomarker profiles. A Bayesian hierarchical model was used to adaptively assign patients to one of four treatment groups using the patient's biomarker profile to estimate the posterior probability of the DCR. The study also had an early stopping rule for futility, in order to drop the potentially inferior treatments from the options available for newly enrolled patients. The study overall had a 46% DCR and identified a higher rate in sorafenib-treated patients with a specific biomarker (KRAS mutations), although sorafenib also resulted in more treatment discontinuations and dose reductions.

Conclusion

Overall there are many potential advantages in using adaptive designs in clinical studies. They can help in earlier selection of the most promising patient characteristics or therapeutic options. If there are mistaken assumptions that have been made prior to the study, adaptive design can be used to correct them midstream. If relevant new data is obtained outside of the trial, an adaptive design can utilize it in the course of the study. It is quite common for the standard of optimal care in treatment of a particular disease to change while a study is in progress. For ethical reasons, and because of the possible effect of the new treatment on the outcomes of the trial, it may be advisable to consider altering the protocol to allow addition of the new treatment in an unbiased way to all of the groups under investigation. Design that makes use of interim data can give investigators new options to modify or re-evaluate the trial while it is underway. Adaptive design can enable investigators to respond to positive or negative surprises from data obtained while the study is in progress. It can be used to stop the study earlier when it becomes clear that there is no benefit to the treatment. In general there are many features of adaptive design that may help to decrease the length of studies and shorten the time of development of investigational drugs. However, it is important to recognize that as part of the design, all potential adaptations that may be undertaken during a trial should be prespecified with the objective of improving the likelihood of a successful and informative trial.

The biggest challenge with these designs is managing the additional logistical complexities and operational details that, without considerable preplanning and careful execution, could impair the validity and integrity of a study. It is extremely important that the interim analyses be done by an independent group (e.g., data monitoring committee) in order to reduce the possibility of introduction of bias. There may be operational biases introduced when using adaptations, and if care is not taken, these may change the trial into a different one that can no longer address the original questions that need to be answered.

RTI (h)(s)™

Health Solutions

References

Bauer P, Köhne K. Evaluation of experiments with adaptive interim analyses. Biometrics. 1994 Dec;50:1029-41.

Branson M, Whitehead J. Estimating a treatment effect in survival studies in which patients switch treatment. Stat Med. 2002 Sep;21:2449-63.

Bretz F, Schmidli H, König F, Racine A, Maurer W. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: general concepts. Biometrical Journal. 2006 Aug;48(4):623-34.

Chow SC, Chang M. Adaptive design methods in clinical trials – a review. Orphanet J Rare Dis. 2008 May;3:11.

Chow SC. Adaptive clinical trial design. Annu Rev Med. 2014;65:405-15.

Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993 Sep;329(14):977-86.

Fox R, Billingham L, Abrams K. Evaluation of methods to adjust for treatment switching in clinical trials. Trials. 2011;12(Suppl 1):A139.

Gallo P, Chuang-Stein C, Dragalin V, et al. Adaptive design in clinical drug development—an executive summary of the PhRMA Working Group (with discussions). J. Biopharm Stat, 2006; 16:275-83.

Garrett-Mayer E. The continual reassessment method for dose-finding studies: a tutorial. Clin Trials. 2006; 3:57-71.

Iasonos A, O'Quigley J. Adaptive dose-finding studies: a review of model-guided phase I clinical trials. J Clin Oncol. 2014 Aug;32:2505-11.

Kim ES, Herbst RS, Wistuba II, Lee JJ, Blumenschein GR Jr, Tsao A, et al. The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov. 2011 Jun;1(1)44-53.

Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. Biometrika. 1976;63:655-60.

Nathan DM, Cleary PA, Backlund JY, Genuth SM, Lachin JM, Orchard TJ, et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. N Engl J Med. 2005 Dec;353(25):2643-53.

Ning J, Huang X. Response-adaptive randomization for clinical trials with adjustment for covariate imbalance. Stat Med. 2010 Jul;29:1761-8.

Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communication in Statistics -Theory and Methods. 1991;20(8):2609-31.

US Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics—draft guidance. February 2010. Available at: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf. Accessed May 3, 2016.