GENERAL STATISTICAL METHODOLOGIES FOR ADAPTIVE DESIGN

Liddy Chen, PhD

PAREXEL
YOUR JOURNEY. OUR MISSION.℠

Background

The Pharmaceutical Research and Manufacturers of America (PhRMA) working group on adaptive design defined adaptive designs as "a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues without undermining the validity and integrity of the trial" While both validity and integrity are important, statistical methods are mainly designed to safeguard validity, which means to provide the correct statistical inferences. E.g., the overall pre-specified Type I error rate must be maintained.

In general, varies statistical methods to control Type I Error in adaptive clinical trials can be classified into the following categories:

1)Repeat testing such as group sequential trials,
2)Multiple hypothesis testing or multiple comparison,
3)Combination of pre- and post- adaptation data.

In more complex adaptive trials where statistical methodologies are not available to ensure the Type I Error rate is controlled at its nominal level, the so called strict control, simulations are usually used to demonstrate the operating characteristic.

Repeat Testing

The earliest development of adaptive clinical trials is perhaps the group sequential trials where a trial can be stopped early if there is strong evidence of efficacy during any planned interim analysis.

- Pocock crossing boundary
- O'Brien and Fleming crossing boundaries
- The alpha spending function developed by Lan and DeMets (1983)

Violation of assumption - the information accumulated each time between interim analyses is statistically independent.

Scharfstein et al. (1997) demonstrated that sequential Wald statistics behave like a standardized partial sum of independent normal variables and thus have independent increments.

Chen et al. (2014) extended the alpha spending function methodology to derive stopping boundaries for testing different hypotheses at different interim analyses.

Combination Test of Pre- and Post- Adaptation Data

The combination methods can be applied to different adaptive designs to provide combined critical values or to combine p-values from pre- and post-adaptation.

Let p_1 denote the p-value based only on the first stage data and p_2 denote the p-value based only on the second stage data.

- If $p_1 \le a$, reject H_0,
- If $p_1 > b$, accept H_0,
- If $a < p_1 \le b$, proceed to the second stage, reject H_0 if the combination function, $C(p_1, p_2) \le c$.

The value c can be solved by the following:

$$a + \int_a^b \int_0^1 \boldsymbol{I}_{[C(x,y)\le c]} dy dx = \alpha,$$

where $\boldsymbol{I}_{[C(x,y)\le c]}$ is an indicator function. Some combination functions are discussed below.

Proschan and Hunsberger (1995) considered a conditional error function,

$$A(z_1) = CP_0(n_2, c|z_1) = \begin{cases} 0, & z_1 < z_b \\ 1 - \Phi\left(\sqrt{z_a^2 - z_1^2}\right), & z_b \le z_1 \le z_a, \\ 1, & z_1 > z_a \end{cases}$$

where z_1 denotes the observed stage 1 test statistic, z_b and z_a are the stage 1 critical value to accept or reject H_0. They need to be selected such that $\int_{-\infty}^{\infty} A(z_1)(z_1) dz_1 = \alpha$, where $\phi(z_1)$ denotes the density of the standard normal distribution. The critical value for stage 2 is a function of stage 1 and stage 2 sample size n_1 and n_2, z_1, and $z_{A(z_1)}$, and can be expressed as $c = \frac{\sqrt{n_1} z_1 + \sqrt{n_2} z_{A(z_1)}}{\sqrt{n_1 + n_2}}$.

Muller and Schafer (2001) also considered a more complex combination rule based on a conditional error function for adaptive group sequential trials.

Brannath's et al (2002) proposed overall p-value for combination tests that was defined by

$$P_C(p_1, p_2) = \begin{cases} p_1 & if\ p_1 \le a\ or\ p_1 > b \\ a + \int_a^b \int_0^1 \mathbf{1}_{[C(x,y)\le C(p_1,p_2)]} dy dx & otherwise \end{cases},$$

where $C(p_1, p_2)$ can be any combination function.

Lehmacher and Wassmer (1999) proposed an inverse normal function:

$C(p_1, p_2) = 1 - \Phi[w_1 \Phi^{-1}(1 - p_1) + w_2 \Phi^{-1}(1 - p_2)]$, with weights $0 < w_i < 1$ and $w_1^2 + w_2^2 = 1$, where Φ and Φ^{-1} denote the standard normal distribution function and the inverse of the function.

Cui et al (1999) developed a down weighting combination strategy for sample size adjustment based on the observed effect size in adaptive group sequential trials. Consider a two-stage design with sample size n_1 and n_2. At the end of stage 1, sample size for stage 2 can be increased to m_2 ($m_2 \ge n_2$) based on observed effect size. Let z_1 and z_2 be the test statistics (realization of Z_1 and Z_2) based on the n_1 and m_2 observation. The test statistics at the end of stage 2 is modified to z^* by keeping the original weights.

$$z^* = z_1 \sqrt{n_1/(n_1+n_2)} + z_2 \sqrt{n_2/(n_1+n_2)}.$$

Bauer and Kieser (1999) proposed a Fisher's combination function, $C(p_1, p_2) = p_1 \times p_2$, for independent samples. When there is no early stopping, simply reject H_0 at the end of the second stage if $p_1 \times p_2 \le c_\alpha = \exp[-\frac{1}{2}\chi_4^2(1-\alpha)]$.

Multiple Hypothesis Testing

The stepwise multiple comparison procedures based on the closure principle, also known as closed testing methods, can also be applied to adaptive clinical trials. These methods are very often used in combination with the stage-wise combination method to allow for additional flexibility. Under the closed testing principal, reject a hypothesis, H_i, if all intersection hypotheses, $H_I = \cap_{i \in I} H_i$, are rejected at level α. For example, the general Bonferroni method, although somewhat conservative, is valid for all correlation structure. A single stage Dunnett procedure was discussed in Bretz's (2006) seamless Phase II/III design.

Bayesian Approach in Adaptive Designs

The Bayes theorem fits naturally in the adaptive design framework (learning and confirming cycle).

- The prior distribution, $P(\theta)$, represents the prior knowledge.
- The posterior distribution, $P(\theta|y)$, represents the knowledge gained after learning from the data, y.

The Type I error or Type II error concept does not exist in the Bayes theorem. These probabilities, or so called operating characteristics, are usually obtained via simulations.

References

- P. Bauer, M. Kieser (1999). Combining different phases in the development of medical treatments within a single trial. *Statistics in Medicine* 18: 1833-1848.
- W. Brannath, M. Posch, P. Bauer (2002). Recursive combination tests. *Journal of the American Statistical Association* 97:236 –244.
- F. Bretz, H. Schmidli, F. Konig, A. Racine, and W. Maurer (2006). Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: General concepts. *Biometrical Journal* 46: 623–634.
- L.M. Chen, J.G. Ibrahim, H. Chu (2014). Flexible stopping boundaries when changing primary endpoints after unblinded interim analyses. *Journal of Biopharmaceutical Statistics* 24: 817-833.
- L. Cui, H.M.J. Hung, S.J. Wang (1999). Modification of Sample Size in Group Sequential Clinical Trials. *Biometrics* 55: 853-857
- K.K.G. Lan, D.L. DeMets (1983). Discrete sequential boundaries for clinical trials. *Biometrika* 70: 659-663.
- W. Lehmacher and G. Wassmer (1999). Adaptive Sample Size Calculations in Group Sequential Trials *Biometrics* 55: 1286-1290.
- H.H. Muller, H. Schafer (2001). Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and classical group sequential approaches. *Biometrics* 57: 886–891.
- M. Proschan and S.A. Hunsberger (1995). Designed Extension of Studies Based on Conditional Power *Biometrics* 51: 1315-1324
- D.O. Scharfstein, A.A. Tsiatis and J.M. Robins (1997). Semiparametric efficiency and its implication on the design and analysis of group-sequential studies. *Journal of the American Statistical Association* 92: 1342-1350.

PAREXEL International
2520 Meridian Parkway
Research Triangle Park
Durham, NC 27713
+1 919 544 3170
www.PAREXEL.COM