

So you want to expand your horizons...

How?

Reminder

- The following techniques seem to be “on-the-fly” trial tinkering
- But! They are carefully thought out and evaluated, and their application anticipated before the trial ever begins
 - Adaptation is not a cure for poor planning
- Regulatory bodies underscore the importance of sound clinical trials practice
 - Usually require reporting at adaptation points

Already Under the Sun...

- Interim analyses are a form of flexible trial design
- Stopping a trial early for efficacy (e.g., α -spending) or futility (e.g., stochastic curtailment) has long been a part of trial design
- Newer techniques allow even more flexibility, though not all potential modifications should or will be acceptable to regulatory bodies

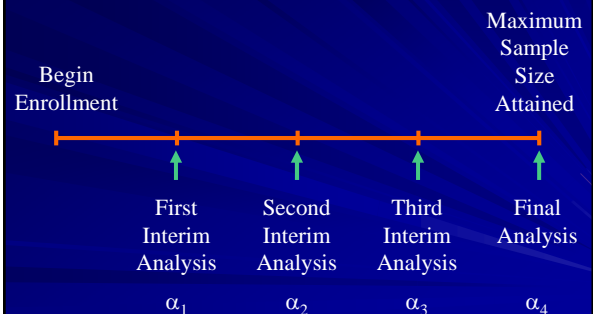
What Can Be Adapted?

- Stopping a trial early
 - Efficacy, futility, harm
- Sample size
- Randomization proportions
- Dosing or duration of treatment
 - Dropping (adding) a treatment arm
- Inclusion / Exclusion criteria
- Primary hypothesis
 - Endpoint
 - Superiority \leftrightarrow Non-inferiority

Stopping for Efficacy

- More familiar to us
- Efficacy: At design stage, determine interim α (significance) levels below which the data would be seen to provide compelling evidence of early treatment success
- In a frequentist trial, the *overall* α level must be “controlled”; classically, it’s set to be less than 0.05

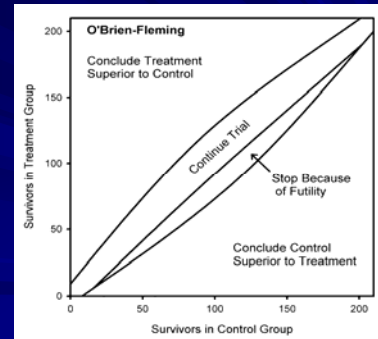
Stopping for Efficacy



Stopping for Futility

- Considers how the treatment appears to be performing thus far, and calculates the likelihood of reaching the predetermined significance level for the trial
- If the likelihood of achieving a final, positive result is too small then the trial is stopped early
- Also known as Stochastic Curtailment / Conditional Power

Stopping Rule at an Interim Analysis



Lewis RJ, Lipsky AM, Berry DA. Clinical Trials. 2007;4:5-14.

Sample Size

- Before a traditional trial is begun, sample size calculations are performed so that adequate power is ensured at the end of a trial
- The calculation requires assumptions about the efficacy of the treatment versus the control, and the within-group variance of the parameter
- What if those assumptions are wrong?
 - e.g., What if the estimated variance you used when planning the trial turns out to be smaller than the trial thus far is demonstrating?

Example Sample Size Calculation

$$2N = \frac{4(Z_{\alpha} + Z_{\beta})^2 \sigma^2}{\delta^2}$$

Sample Size Re-estimation

- At an adaptation point it might become apparent that the expected variance (or effect size) is not what was used in the sample size calculation
- Or more generally, that the information hoped to be gained from the trial will require more (or fewer) patients than initially planned
- Adaptation: Re-estimating the required sample size, thus (for instance) continuing enrollment past the end of the initially proposed trial size

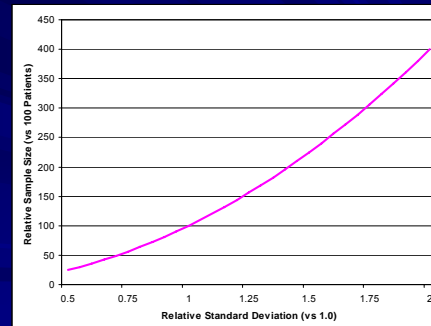
Sample Size Re-estimation

- The decision to extend the trial should represent an earnest reassessment of the variance (and/or the definition of a clinically important treatment effect)
- Note that external information may also be used
 - e.g., Using the variance from a similar population in a recently completed study
 - Though this makes modeling rather difficult

Sample Size Re-estimation

- Importantly, re-estimation is not always about increasing the size of the trial
- If the variance is less than anticipated, then you may end up enrolling *too many* patients in the trial
- If the trial enrolls the initially proposed number of subjects, you may be harming those patients in the poorer performing arm
 - e.g., By continuing to enroll patients into the placebo arm

Sample Size Re-estimation



Randomization Proportions

- Trials are a compromise between treating patients effectively and maximizing information gained
- From a dispassionate scientific perspective, maximum information or power is gained (ignoring prior information) by having equal numbers of subjects in the treatment and control arms (1:1)
- However, as evidence accumulates that one arm is performing better, we should feel compelled to preferentially assign patients to the better performing arm

Randomization Proportions

- For instance, a trial may begin with 1:1 assignment of subjects, but then switch to 1:2 or 1:3 (etc.) assignment ratios if it turns out that the subjects in the second arm have better outcomes
- In addition to this “response adaptive” system, there are also “treatment adaptive” and “covariate adaptive” systems which allow for continued rebalancing of groups to ensure the desired proportions of subjects (and covariates) in the arms

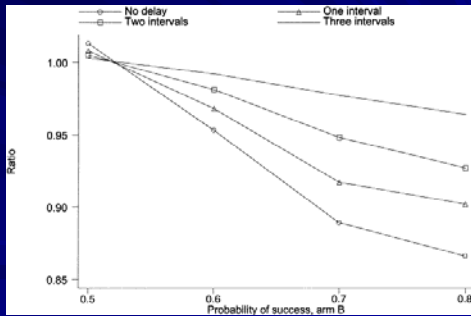
Treatment Adaptive

- In a standard trial with random assignment of treatment allocation, it is possible for there to be an imbalance in the distribution of treatment vs control solely due to random chance
 - For instance, among the first 100 patients, 57 may have been assigned to treatment and 43 to placebo
- If a treatment adaptive design were planned, then the subsequent patients may be randomized at, say, a 1:2 odds of being in the treatment group until the groups were rebalanced
 - The assignment remains random, just skewed

Covariate Adaptive

- Similarly, if, in your study, it seemed that an important covariate were over-represented in one group, you may want to have planned to use covariate adaptive randomization to rebalance the groups
 - Consider a head trauma study with an imbalance in “GCS on arrival”
 - While you can always control for this confounder at the end of the analysis – validity concerns aside – there are fewer patients in one of the arms, and thus a higher variance

Randomization Proportions



Karrison TG, Huo D, Chappell R. *Controlled Clinical Trials*. 2003;24:506-522.

Adaptive Dosing

- Adaptation can be used to add or drop dosing regimens (arms) while a trial is in progress
- In Phase II trials, available dosing regimens are traditionally decided a priori; with adaptation, we can efficiently and in real-time allow the trial to tune itself to the best performing dose – even doses not initially part of the trial
 - Often used in oncology

Adaptive Dosing

- If a dose is underperforming (or harmful) based on pre-established criteria, that dose level is excluded from further patient enrollment
 - We randomize zero patients to that arm
- If it seems the most effective dose lies between two doses being used, we can allocate patients to a new “in-between” arm (even for a continuum of doses)
- In addition to efficacy, preferential allocation to specific doses (as with randomization) addresses the need to increase precision in those arms

Adaptive Dosing

- To clarify that last point:
 - First, identify the doses that appear to be most effective (and least harmful)
 - Then assign sufficient numbers of patients to those remaining doses to increase our information about them (i.e., increase the precision of the estimates)
- This technique was used in a large stroke trial investigating a neuroprotective agent (2003)

Adaptive Dosing

Acute Stroke Therapy by Inhibition of Neutrophils (ASTIN) An Adaptive Dose-Response Study of UK-279,276 in Acute Ischemic Stroke

Michael Krams, MD; Kennedy R. Lees, MD; Werner Hacke, MD; Andrew P. Grieve, PhD;
Jean-Marc Orgogozo, MD; Gary A. Ford, MD; for the ASTIN Study Investigators

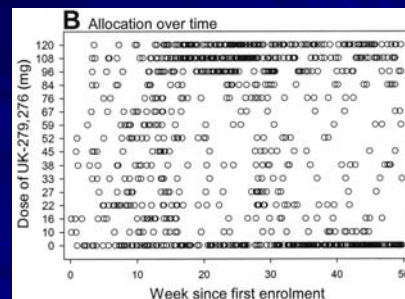
Background and Purpose—UK-279,276 (neutrophil inhibitory factor) reduced infarct volume in a rat middle cerebral artery occlusion reperfusion model. ASTIN (Acute Stroke Therapy by Inhibition of Neutrophils) was an **adaptive phase 2 dose-response-finding**, proof-of-concept study to establish whether UK-279,276 improves recovery in acute ischemic stroke. The prime objective was to **determine the dose that gave a clinically relevant effect** in patients.

Methods—A Bayesian sequential design with real-time efficacy data capture and continuous reassessment of the dose response allowed **double-blind, randomized, adaptive allocation** to 1 of 15 doses (dose range, 10 to 120 mg) or placebo and **early termination for efficacy or futility**. The primary end point was change from baseline to day 90 on the Scandinavian Stroke Scale (SSS), adjusted for baseline SSS, aiming for a 3-point additional mean recovery above placebo.

Results—Nine hundred sixty-six acute stroke patients (887 ischemic, 204 cotreated with intravenous tissue plasminogen activator; mean baseline SSS score, 28; range, 10 to 40) were treated within 6 hours of symptom onset. Mean Δ SSS was approximately +17 points of improvement on SSS for the overall evaluable population. There was no treatment effect for UK-279,276 (posterior probability of futility, 0.89). The trial was stopped early for futility. Post hoc analysis indicated a mean 1.6-point additional improvement on Δ SSS in the tissue plasminogen activator-treated subset (credible interval=0.5, 2.6). UK-279,276 was generally well tolerated, with no increased incidence of infections.

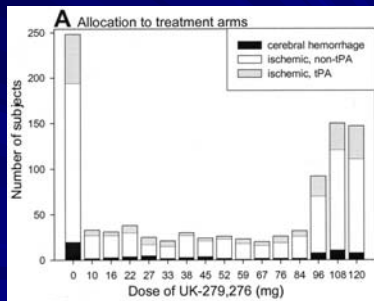
Conclusions—UK-279,276 did not improve recovery in acute ischemic stroke patients but was devoid of serious side effects. **The adaptive design facilitated early termination for futility.** (*Stroke*. 2003;34:2543-2548.)

Adaptive Dosing



Krams M, Lees KR, Hacke W, et al. *Stroke*. 2003;34:2543-2548.

Adaptive Dosing



Krams M, Lees KR, Hacke W, et al. Stroke. 2003;34:2543-2548.

Other Adaptive Schemes

- Inclusion / Exclusion criteria
 - An inclusion or exclusion criterion may be reducing the rate of enrollment
- Primary endpoint
 - We have several possible endpoints of interest, all of which were identified at the beginning
 - We may, perhaps, choose one as the trial proceeds based on least amount of missing data or smallest variance
- Superiority ↔ Non-inferiority
 - Changing the goal (hypothesis) of the trial

Caveat

- There are some concerns, especially with changing the inclusion/exclusion criteria
- Trials are microcosms of the greater population of patients who may benefit from them
- The results of trials are extrapolated to populations represented in those trials
- If the population from whom the trial draws continually changes, it may be difficult to determine to which population the results apply
 - i.e., There is an external validity concern

More on Adaptation Points

- Regulatory agencies are concerned (as you should be) that leakage of the interim results may bias future enrollments or treatment assignments, or in some other way bias the trial
- They request the utmost confidentiality surrounding the interim assessments of the data
 - Need-to-know basis is important
- This is not obviated via a Bayesian approach

Verifying Performance

- Simulations
- Reality checks

Simulations

- Standard trials are usually amenable to clean solutions – the math has been around for decades and look-up tables abound
- With the addition of adaptation, the solutions are a lot less trivial
 - Solutions: Determining the expected overall type I and II error, and getting a handle on the expected sample size, all of which are of paramount importance to the regulators
- How do we accomplish this?

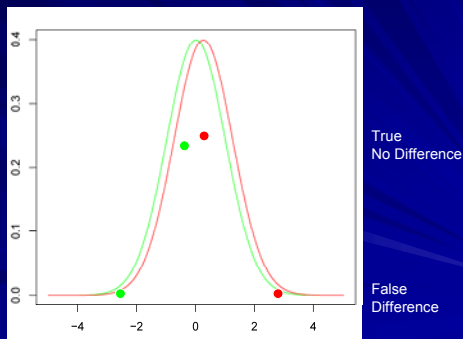
Trial Simulation

- The characteristics of the trial are encoded (e.g., group sequential, 3 blocks of 8 subjects, with response-adaptive randomization), and virtual subjects are enrolled
- Enrolling subjects in a single trial will yield a specific (virtual) study result
- Repeating this many times will generate a distribution of study results

Trial Simulation

- Because we control the distributions from which the computer randomly samples the outcomes for each subject, we can calculate frequentist error rates
 - To evaluate expected Type I error, we specify that the treatment and control effects are equal (H_0)
 - As the computer randomly samples from these distributions, some trials will conclude that the treatment arm is better than the control arm
 - The fraction that these trials represent is the Type I error rate

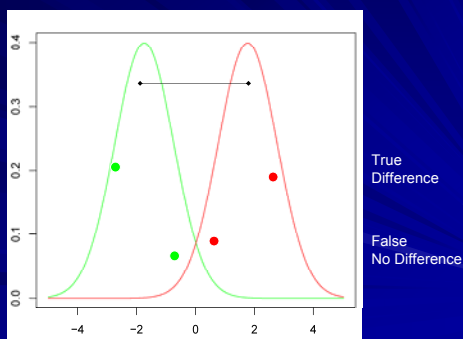
Simulation Under H_0



Trial Simulation

- Similarly, we can specify that the difference in outcomes is equal to the treatment effect we wish to be able to detect
- The trials that conclude equivalence represent the fraction of trials with type II error ($= 1 - \text{Power}$)
- The trial can also be run using what you believe about the distributions (i.e., by placing priors on them) to compute an expected sample size under those assumptions
- As the trial progresses the computer will preferentially assign virtual subjects via the pre-specified response adaptive randomization scheme

Simulation Under H_A



Reality Checks

- Consider carefully where your planned trial may take you
 - Simulation results, especially the outliers, are useful in examining potentially unexpected consequences of the trial parameters
- Examples:
 - Too few patients in the control arm (ECMO)
 - Specify a minimum desired
 - The trial may never stop
 - Specify a maximum desired

A Bayesian Approach

- From a frequentist (standard statistical) perspective, the design and the results of a trial are an inseparable package
 - This complicates our ability to learn from the data, or to use the data for flexible decision making
- A Bayesian approach is able to isolate the data from the design thus overcoming these limitations
 - A “downside” is the need to specify prior information, potentially in the form of subjective beliefs
- Because the data are separate from the design, Bayesian-designed trials are especially amenable to adaptation (“flexible design”)

A Bayesian Approach

- We are working on designing two-armed trials (e.g., treatment / control) for which there is a binary outcome (e.g., survival / death)
 - Rapid, binary outcomes are especially useful in ED research
- The trials minimize the “costs” of enrolling patients; of subjects suffering the worse of the two possible outcomes; and of making the wrong decision at the end of the trial (analogous to frequentist type I and II error)
- The parameters we specify as the investigators at the beginning of the design-phase allow us to explicitly balance these three competing priorities

A Bayesian Approach

- E.g., If a large proportion of those suffering from a disease are enrolled in the trial, we may consider survival of our subjects more important than (though not to the point of disregarding) minimizing erroneous scientific conclusions
 - This is similar to response adaptation, here balanced against the risk of “scientific” error, though its use is to minimize the overall cost
- The designs may choose from one of three randomization ratios as each new group of patients is enrolled
 - Adaptive randomization
 - Helps us more efficiently achieve the goals stated above

A Bayesian Approach

- At each of one of the prespecified adaptive points, early stopping or changing randomization ratios is permitted
- In the trials illustrated on the next slide, there were 3 blocks of 8 patients
 - i.e., 8 patients were enrolled, then a decision was made to either stop or continue the trial; this was potentially repeated two more times
 - The unit of cost is equal to the enrollment of a single patient
 - Adaptive randomization and/or adding a cost for “failure” was permitted in some designs

Trial Characteristics

Adaptive	Random-ization	Add'l Cost of Failure	Exp. Sample Size	Exp. Cost	Type I Error Rate	Type II Error Rate
No	1:1	0	26.2	394	0.04	0.22
Yes	1:3, 1:1, 3:1	0	23.5	175	0.05	0.21
		10	23.0	(691)	0.05	0.20

Example Differences

Control Survival	Test Survival	Optimal Action Standard Failure = 0	Optimal Action Adaptive Failure = 0	Optimal Action Adaptive Failure = 10
0 of 7	6 of 9	Continue (1:1)	Continue (3:1)	Stop; Test Better
2 of 7	4 of 9	Continue (1:1)	Continue (1:1)	Stop; No Difference

Online Tools

- MD Anderson
 - <http://biostatistics.mdanderson.org/SoftwareDownload/>
 - Lots of good utilities, including “Adaptive Randomization” to help with response adaptive trials
 - Allows 10 arms; minimum number of patients before adapting randomization scheme; maximum number of patients or length of trial
 - Free
- Cytel
 - <http://www.cytel.com/>
 - East & EastAdapt
 - Lots of control, including sample size re-estimation and preservation of type I error
 - Not Free

Conclusions

- Not all trials need (or should have) adaptive designs
- When used appropriately and conscientiously, adaptive designs, compared to standard designs, may maximize the information we obtain from the trial and/or minimize the harm we cause subjects
 - An adaptive design will not save a poorly planned trial
- The next step is designing adaptive trials from a Bayesian perspective, though the regulatory bodies are only now beginning to embrace this methodology

Questions?

The Englishman's strong point is his vigorous insularity; that of the American his power of adaptation. Each of these attitudes has its perils. The Englishman stands firmly on his feet, but he who merely does this never advances. The American's disposition is to step forward even at the risk of a fall.

Thomas Wentworth Higginson
1823–1911

References

- Chow S-C, Chang M. Adaptive Design Methods in Clinical Trials. Chapman & Hall / CRC; 2006.
 - An introduction geared toward statisticians
- Berry DA. Bayesian clinical trials. Nat Rev Drug Discov. 2006;5:27-36.
 - Review of Bayesian methodology, as well as their extensions into adaptive design