

Late 'Learn Phase' Designs

Valerie Durkalski-Mauldin, MPH, PhD
durkalsv@musc.edu

Clinical Development

- Phase I- “Dose Finding”
 - Often focus on safety/tolerability
- Phase II – “Safety and Indication of Efficacy”
 - Tolerability
 - Feasibility
 - ‘Should we move forward with this specific treatment....Phase III Go/No-Go decision’
- Phase III-“Confirmatory”
 - Definitive evidence of efficacy
- Phase IV- “Post Marketing Surveillance”
 - Rare side effects

Past Collaborators.....



CTMC Aug 7-11, 2017

Phase III

- Generally require 'larger' sample sizes (larger budget).
- Promising agents in preclinical setting and pilot/early studies do not guarantee confirmation of efficacy.
- Need to start 'getting it right' early on.
 - If we know several will fail then 'fail fast and fail often'...and make this information public (publish).

Why early studies seem promising.....

- Very selective study population
- Small number of sites
- Use of historical controls
- Use of surrogate endpoints

Phase II Goals

Overall Goal: Should one proceed to a phase III trial to confirm efficacy?

- Estimate the frequency and severity of side effects (tolerability)
- Exploratory for efficacy.... identify treatments with potential efficacy (surrogate or true endpoints)
- Quickly discard treatments without promise (futility)
- How feasible is the study protocol in terms of: compliance, administration, delivery, and cost? Can patients be enrolled?

Advantages of Phase II trials

A futility analysis can be built into your Ph3 design, but there is a lot to learn during phase 2.

- Gives investigators experience in trial procedures
- Points out needed future protocol modifications (feasibility, nonadherence, recruitment)
- Gain experience with implementation in multiple sites
- Can provide insight on blinding procedures
- Helps in planning resources for Phase III

Phase II Designs

- Traditional designs to show an indication of efficacy
 - Fixed sample single stage designs
 - Two-Stage Designs (Gehan, 1961; Simon, 1989, Fleming, 1982)
- Futility Designs (Palesch 2005; Tilley, 2006; Yeatts 2013)
- Randomized Selection Designs (Simon, 1985)
- Hybrid Designs, Seamless Designs.....

Phase II Design Considerations

- Single or Multi-Arm: can historical controls be used; can multiple treatments/doses be included
- Fixed-sample trial: the number of patients allocated to the one (or more) treatments is fixed before the study begins.
- Sequential/multi-stage trials: the decision whether to continue taking new patients is determined by the results accumulated to that time.

Single-arm, Single Stage Design

- Clinically important response rates (e.g., *higher is better*):
Standard response rate (historical control), p_o
Target response rate, $p_T = p_o + a$
- $H_o: p_T \leq p_o + a$ vs $H_A: p_T > p_o + a$
Reject null => Preliminary evidence of efficacy.
Fail to Reject => No evidence of efficacy.
- Type I and II error rates (α and β)

*Two-arm version—beware of the underpowered Phase III trial

Optimal Two-Stage Designs (Simon 1989)

- To minimize patients exposed to ineffective treatments
- Stage 1/Stage 2 (single arm)
 - Enroll n_1 patients into Stage 1 (where n_1 is small).
 - If number of responses $\leq r_1$, stop at Stage 1 b/c the treatment appears ineffective.
 - Else enroll n_2 patients in Stage 2.
- Extensions to this method (Shuster, 2002)

Limitations to the classic single arm Phase II trial design

- evaluating each experimental regimen/agent individually is limited
- Difficult to separate trial effects (e.g., patient selection, trial eligibility, imaging techniques, assessment schedule, treatment locations) from treatment effect on clinical outcomes.

Futility Designs

Futility Designs

- Goal: designed to identify which treatments are least likely to demonstrate benefit
- Features:
 - One-sided hypothesis
 - Formulation of the typical null and alternative hypothesis are reversed (want to prove futility)
 - Null: treatment has promise
 - Can be one-sample (single arm) or two-sample

Futility Design (parameters)

- Standard/Placebo response (success) rate = p_o
- Δ = the improvement considered clinically meaningful and not futile
- Target response rate: $p_T = p_o + \Delta$
 - Minimum proportion of successes in the treated group to consider the treatment for further study.

One-Sided Futility Hypotheses:

$$H_o: p_T \geq p_o + \Delta$$

$$H_A: p_T < p_o + \Delta$$

- Reject null => Futile to study further
- Fail to Reject => No evidence of futility, Proceed to Phase III to confirm efficacy

Interpretation of Errors

	Type I error α	Type II error β
Traditional Efficacy Trial	Ineffective therapy is brought forward (false positive)	Effective therapy is missed (false negative)
Phase II Futility Study	Effective therapy is missed (false negative)	Ineffective therapy is brought forward (false positive)

Single Arm

Futility Design Example

A study is going to test the futility of treating Parkinson patients with an experimental drug. The goal is to reduce the number of patients who require levodopa (treatment failures).

Historical data indicates that untreated patients would have a failure proportion of 42.5% (p_o) after 12 months of followup.

Investigators defined the target treatment response rate as 32.5% (p_T); $\Delta=0.10$.

$$H_0: p_T \leq (p_o - 0.10) \text{ versus } H_A: p_T > (p_o - 0.10)$$

One-sided alpha = 0.10 (error rate for falsely claiming futility)

Beta = 0.15 (85% power; 15% chance of moving forward with an ineffective treatment)

Single Arm Example (cont)

- **Ho:** $p_T \leq (p_o - 0.10)$ versus **H_A:** $p_T > (p_o - 0.10)$
- *One-sided test for a binomial proportion:*
 - *Observed $p_T = 42\%$ ($p_o = 42.5\%$)*
 - *pvalue < 0.01*
 - *Reject the null and claim futility*

Single Arm Example (cont)

- **Ho:** $p_T \leq (p_o - 0.10)$ versus **H_A:** $p_T > (p_o - 0.10)$
- *One-sided test for a binomial proportion:*
 - Observed $p_T = 14\%$ ($p_o = 42.5\%$)
 - *Pvalue* > 0.99
 - *Fail to reject the null **BUT** also cannot confirm efficacy*
- **Futility designs are not designed to prove efficacy (different error rates, specific population, few sites.....)**

Two Arm Futility Design Example: I-DEF (NCT01662895)

- Phase I CRM design identified maximum tolerable dose (*Salim, Yeatts et al Stroke 2011*).
 - Based on safety; need information on efficacy
- Testing the futility of treating ICH patients with an experimental drug (DFO).
 - The goal is to increase the number of patients who have a 90-day mRS of 0-2 (treatment success).
- Historical data in this population is potentially unreliable (temporal changes in patient care; protocol adherence; variability across centers).

Two Arm Futility Design Example: I-DEF (NCT01662895)

- Investigators anticipate placebo arm success to be 28% (p_o).
- Defined the minimum clinically important difference to be 12% (Δ).

$$\mathbf{H_0: } p_T \geq (p_o + 0.12) \text{ versus } \mathbf{H_A: } p_T < (p_o + 0.12)$$

One-sided alpha = 0.10 (error rate for falsely claiming futility)

Beta = 0.20 (80% power; 20% chance of moving forward with an ineffective treatment)

Futility Designs

- Design useful for identifying ineffective therapies
- Carefully consider choice of your error rates as well as the threshold for futility
- Can be used in combination with other designs

Selection Designs and Hybrids

Randomized Selection (Phase II) designs

- Typically used to select one best treatment/dose of many in the early phase
- Multi-arms with or without control
- No formal hypothesis testing (type I error under equal efficacy assumption is not of interest)
- Pre-specified selection procedures
 - Want to have a high probability of correctly selecting the superior treatment/dose

Simple Selection Design (Simon et al, 1985)

- Simple idea of selecting the best among the K trts/doses to take forward to Phase III based on the statistical ranking and selection theory (Gibbons et al, 1977).
- Sample size estimated to ensure that if the best treatment is superior by at least D , then it will be selected with high probability (e.g., 90%).

Simple Selection Design (cont'd)

- The design will always select the “best” treatment, but if the difference between the “best” and the second best is $< D$, the probability of correct selection will be $< 90\%$.
- Somewhat arbitrary and no information that even the “best” treatment will be worth proceeding to a Phase III.
- Most trials desire a quantification of the difference between two treatments, not just the selection of the best one (extensions to these methods).

Hybrid Designs

- Remaining question on most efficacious dose
- Eliminate inferior treatments quickly so that more resources can be allocated to the likely winner(s)
- Two-stage design: Selection + Futility
 - Example in ALS Trial of selecting one of two doses of CoQ10 (Levy et al, 2006).

Selection + Futility (Levy et al, 2006)

- Combination of the Simple Selection (Stage 1) design that includes a control as one of the K treatments and a concurrently-controlled Futility (Stage 2) design.
- ‘Seamless’ Learning Design
 - data from both stages are utilized in Stage 2 for the comparison between the selected treatment and control.

Selection + Futility (Levy et al, 2006)

- Control plus two doses of active treatment (CoQ10)
 - 35 patients per arm (Stage 1 Total N=105)
 - 40 patients per arm (Stage 2 Total N=80)
 - Final analysis has 75 patients per group (active/control)
- Primary outcome: Change from baseline in 9 mos functional rating scale (ALSFRS_r)

Stage 1 - Selection

- Choose one of two doses based on smaller mean change in outcome (smaller decline).
- Designed to have a probability of 80% for selecting the correct dose.
- If doses are equal then dose determined by safety and cost.

Stage 2 - Futility

- Randomize to chosen dose or placebo (n=40)
- Investigators anticipate placebo arm mean decline to be 8.5% (SD 8.4).
- Defined the minimum clinically important difference to be 20% (Δ).

H₀: $(p_T - p_o) / p_o \geq 0.20$ versus H_A: $(p_T - p_o) / p_o < 0.20$

One-sided alpha = 0.10

Beta = 0.20 (80% power)

Multi-Stage/Hybrid designs

- When do we use these designs?
 - Based on the questions needing to be answered
 - Based on clinical setting
- Improved efficiency in terms of sample size and administration/infrastructure.
 - Conducting one trial rather than two
- Increased complexities that require strong infrastructure and planning.....But this should not be seen as a deterrent.

Choosing a Phase II Design

- Phase II Design choice depends on:
 - Outcome of interest
 - Indication of safety, feasibility, efficacy or futility
 - Availability of surrogate outcomes
 - Inclusion of control group
 - Multi-dose/trt (select 1 to move forward)
 - Single Stage or Multi-stage (early stopping)
 - Rate of Recruitment and Length of follow-up

Select References

- Palesch Y, Tilley B, Sackett D, et al. Applying a Phase II Futility Study Design to Therapeutic Stroke Trials. *Stroke* 2005; 36: 2410-2414.
- Tilley B, Palesch YY, Kieburtz K et al. Optimizing the ongoing search for new treatments for Parkinson disease: Using futility designs. *Neurology* 2006; 66: 628-633.
- Yeatts S, Palesch Y, Moy C, Selim M. High Dose Deferoxamine in ICH (HI-DEF) Trial: Rationale, Design and Methods. *Neurocrit Care* 2013; 19: 257-266.
- Levy G, Kaufmann P, Buchsbaum R, et al. A two-stage design for a phase II clinical trial of coenzyme Q10 in ALS. *Neurology* 2006;66:660-3.
- Simon R, Wittes RE, Ellenberg SS. Randomized phase II clinical trials. *Cancer Treatment Reports* 1985;69:1375-81.