VALOR
A Case Study of a Trial with an SSR Including a Promising Zone

Zoran Antonijevic – Cytel Inc.

Introduction

Presentation Overview

- Introduction
- The VALOR Trial
- Implementation/Challenges

Benefits from Adaptive Methods

- Reduce costs, time of development
  - Combining stages of development, early stopping for efficacy/futility
- Adaptive approach flexibility improves dose and sub-population selection
  - Flexible designs at exploratory stage of development
  - Two stage confirmatory designs with dose or sub-population selection
- **De-risk investments at late stage of development**
  - **Unblinded SSR (Promising Zone)**, GSD

Challenges with Adaptive Design

Potential for operational bias:

- Inadvertent (or intentional?) unblinding of interim results
- Investigator behavior changes after interim

Design itself can reveal the information about interim data

- Reports by financial analysts

Inflation of type I error rate

- Handled by appropriate statistical methodology

Practical Considerations

- Things may not go as planned

The VALOR Trial

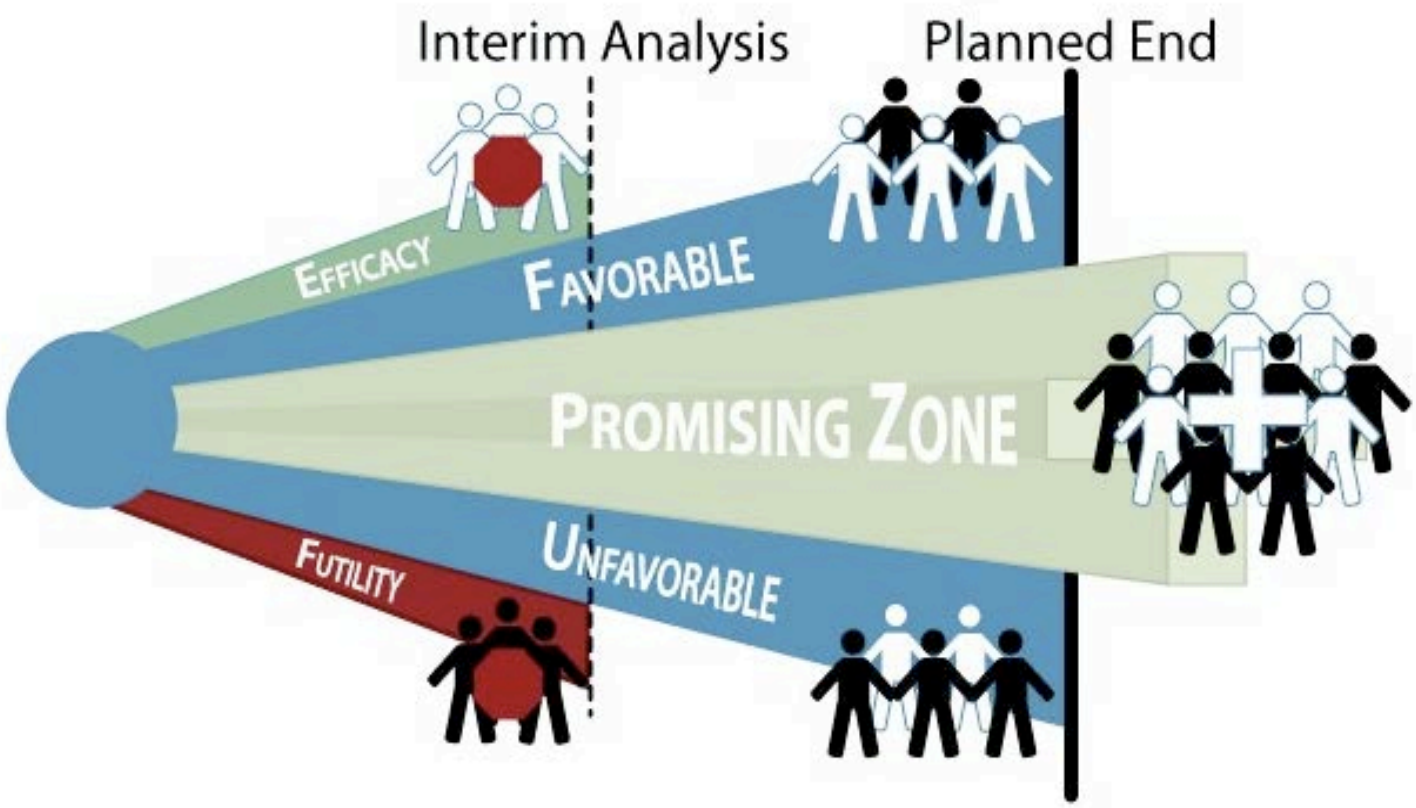
Case Study: VALOR Trial for AML

- Therapy for relapsed or refractory AML is generally unsatisfactory; no approved drugs; dismal prognosis
- Phase 3, double-blind, placebo-controlled, multinational trial for first-relapsed or refractory Acute Myeloid Leukemia (AML)
- Evaluate efficacy (PE = Overal Survival) and safety of Vosaroxin+Ara-C versus Ara-C+Placebo

Sponsor's Dilemma

- Based on phase 2 data:
  - Assume 5/7 month median for Ctrl/Trtm (HR=0.71)
  - Require 375 events and 450 subjects @ 19/month
- But phase 2 estimates are subject to uncertainty
  - What if 5/6.5 mth median on Ctrtl/Trtm (HR=0.77)?
  - HR=0.77 is still clinically meaningful
  - Require 616 events and 732 subjects @ 31/month
  - Not a feasible option for sponsor
- Given these constraints, how to design this single pivotal trial?

Study Design Chart



Efficacy zone (OBF)
One-sided p=0.0015

Favorable zone
(CP ≥ 0.9).

Promising zone
(0.3 ≤ CP < 0.9);

Unfavorable zone
(0.1 <CP < 0.3);

Futility zone
(CP ≤ 0.X)

Interim Analysis at 187 Events
Planned End at 375 events
Maximum number of Events: 561

VALOR Results at Interim Analysis

- Interim was conducted at 173 events, rather than 187 as planned
- HR was 0.76
- Conditional Power was 82%, in the promising zone, so sample size was increased
- Both sample size and events were increased by 50%

Interim Result Triggered Additional Investment

Milestone-Driven Investment: Sunesis Pharmaceuticals to Implement One-Time Sample Size Increase to Phase 3 VALOR Trial in AML. DSMB Recommends Increase Following Single, Pre-Planned Interim Efficacy and Safety Analysis of VALOR; Enrollment Completion Expected in 2013. DSMB Recommendation Triggers $25.0 Million Investment in Sunesis from Royalty Pharma.

Press Release, September 11, 2012. Sunesis Pharma, South San Francisco

Final VALOR Results

- Primary Endpoint *Overall Survival*:
  - unstratified results:
  - stratified results: HR = 0.83, p=0.02
- Single secondary endpoint, *Complete Response rate*: 30.1% vosaroxin arm vs. 16.3% placebo arm, p<0.0001.

Final VALOR Results

- Primary Endpoint *Overall Survival*:
  - 7.5 months on Vos vs. 6.1 months on Placebo.
  - unstratified results: HR = 0.87, p=0.06
  - stratified results: HR = 0.83, p=0.02
- Single secondary endpoint, *Complete Response rate*: 30.1% vosaroxin arm vs. 16.3% placebo arm, p<0.0001.

Implementation/ Challenges

Challenges Revisited

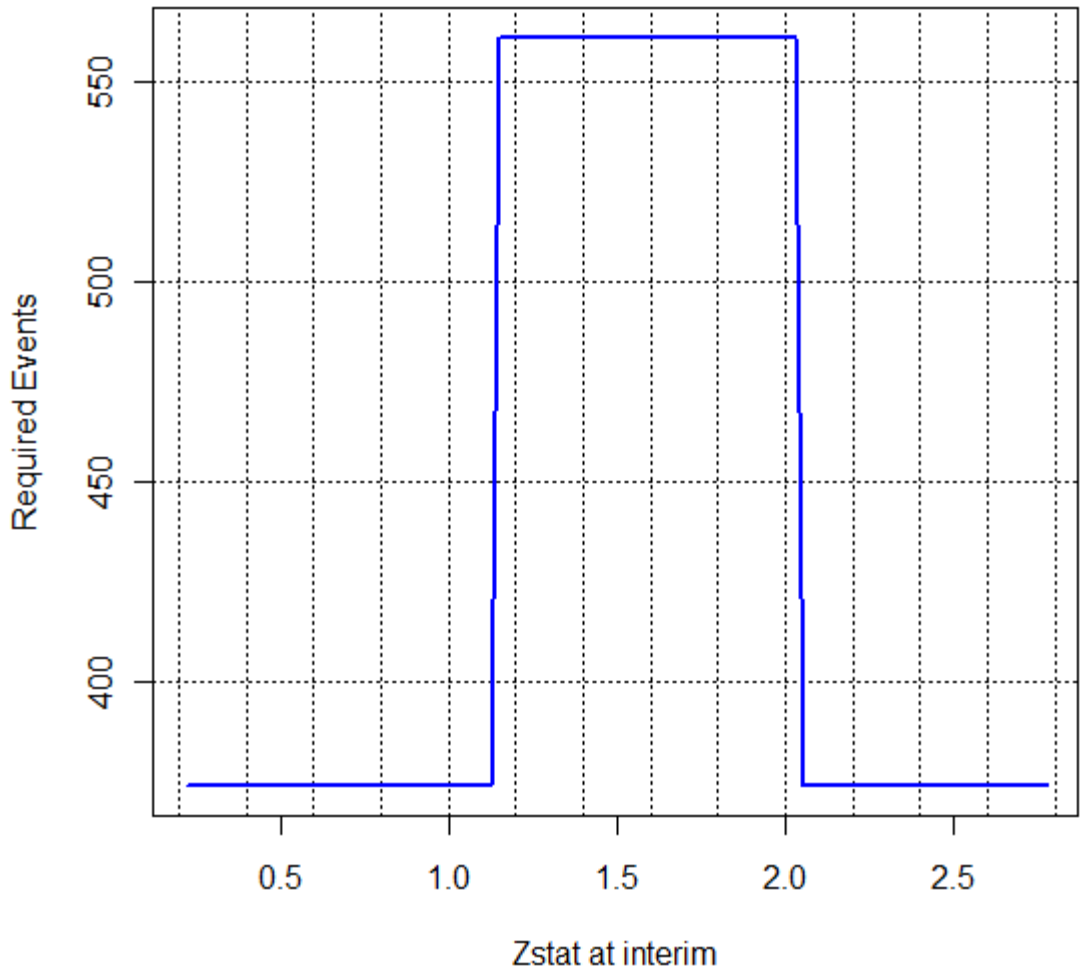
- Potential for operational bias
- Design itself can reveal the information about interim data
- Inflation of type I error rate
- Practical Considerations

Operational Bias

Implementation of Adaptive Designs most successful with:

- Objective endpoint measures (overall survival)
- Double-blind study
- Limited knowledge of rules for interim analysis
  - Sunesis personnel not aware of specific rules for DMC
- Limited knowledge of interim results
  - One-time fixed sample size increase (limit back calculation, to be revisited)
  - Consider communication plan
  - Consider using a system to control the information flow. In this case Cytel's ACES was used.

Design Revealing the Interim Data



Design Revealing the Interim Data

- With unblinded SSR step increase preferred over the linear increase from implementation standpoint
  - Loss in efficiency relatively minor
- Cytel had only a few examples where a sponsor chose the linear increase, but that complicated implementation.
  - In this situation sponsor and study team are not informed of the new sample size after the interim
  - The unblinded statistician given the task to monitor enrolment/accrual of events and inform about upcoming enrolment closure.
- In some situations an increase in sample size “by blocks” recommended.

Preserving the Type-1 Error

- Let D_1 and D_2 be the pre-specified total events at interim and final analysis. (Here $D_1 = 187$ and $D_2 = 375$)
- Let LR_1 and LR_2 be the corresponding logrank statistics
- Suppose D_2 is altered to $D_2^* > D_2$ at the interim
- Let LR_2^* denote the corresponding altered logrank statistic
- Type-1 error is preserved if we use

$$Z_{\mathsf{chw}} = \sqrt{\frac{D_1}{D_2}} \times \mathsf{LR}_1 + \sqrt{\frac{D_2 - D_1}{D_2}} \times \frac{\sqrt{D_2^*}\mathsf{LR}_2^* - \sqrt{D_1}\mathsf{LR}_1}{\sqrt{D_2^* - D_1}}$$

instead of LR_2^* for the final analysis

CHW Adjustment, cont'd

- The interim happen at 175 events/375 instead of 187
  - CHW adjustment: weights for increments remain as planned (187/375) but the calculation of the increments will change.
  - O'Brien-Fleming: alpha spend based on the observed information fraction 175/375.
  - Event prediction tools very useful

Practical Considerations

- Data change after the interim…
  - Often data are not perfectly clean at the interim
  - Should the observed interim test statistic on which sample size adjustment was made be used?
  - Or should the CHW statistic be recalculated based on final data?
    - If based on final data, then how is the cutoff redefined?
- Cannot change SAP once interim is passed

Concluding Remarks

- Regulatory Assessment
  - The FDA requires another trial to demonstrate efficacy
  - The EMA gave a nod to submit a marketing application
- Adaptive design played an important role as it allowed staged investment
- Careful implementation led to
  - Control of Type I error
  - Minimized potential for operational bias
- Practical questions emerge with implementation; we need to learn from experiences
- Beware of uncertainty
  - Time to event endpoints particularly tricky

References

- Mehta, C.R., and Pocock, S.J. (2011). Adaptive increase in sample size when interim results are promising: A practical guide with examples. *Statistics in Medicine*, 30(28):3267-3284.
- Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics.* 1999 Sep;55(3):853-7.
- Wassmer G. Planning and analyzing adaptive group sequential trials. *Biom J.* 2006 Aug;48(4):714-29.