New and Innovative Design Approaches for Phase III Oncology Trials, with Emphasis on Regulatory Expectations

- Jonathan R. Smith, Ph.D.

September 11th 2007, Philadelphia

Pharma Ed, 1st Adaptive Trial Designs conference

Outline of Presentation

- Introduction
- Traditional Interim Analysis for Efficacy or Futility
- Adaptive Re-Sizing
- Phase IIB/III Designs Incorporating Treatment Selection & Testing
- Phase III designs with Safety-based Treatment Selection
- Phase IIB/III Designs with Adaptive Selection of Patient Population
- Phase II/III Designs for both Accelerated Approval & Full Approval
- Regulatory Considerations for Adaptive Designs
- Concluding Comments

Need for Improved & Innovative Designs in Cancer Drug Development

- **400+ New Molecules for Cancer** currently in clinical development [PhRMA data]
- **81 Months Average Time** from IND to NDA for Anti-Neoplastics [Tufts, based on 1996-1998]
- **Oncology Drug's Failure Rate in Phase III is 59%** [Kola & Landis, 2004]
  - CMR even quotes a **71% Oncology Failure Rate in Phase III** [based on # enter Ph III/# get to mkt - see Peck (2003)]

Emphasis Within this Presentation

- Focus of this presentation on Late Stage Parallel Group Oncology Trials:
  - Phase III alone OR
  - Phase II/III Combined
- Primary Endpoint under consideration is generally Time-to-Event, e.g.,
  - Overall Survival (OS), Progression-Free Survival (PFS), or Disease-Free Survival (DFS)

Interim Analysis for Futility in Pivotal Oncology Trials

- Futility is Particularly Important for Ethical Reasons in Ph II/III Oncology Trials and in Ph III when Ph II was single arm
- Practicality of Stopping for Futility in Oncology with Time-to-Event Primary Endpts is often an issue
  - Generally need 40%-50% of target # of events
  - Enrollment often over by that stage
  - Cost Savings and unnecessary further exposure of pts can still be obtained by stopping study after enrollment has been completed,
  - Alternatively sometimes useful to base futility assessment on Intermediate Endpt (IE) that is Predictive of Clinical Endpt (CE)

Interim Analysis for Futility in Pivotal Oncology Trials, cont'd

- Choice of Method for Futility Assessment is critical:
  - Should have Minimal Negative Impact on Power if compound is truly efficacious
  - Traditional Conditional Power based approach at early look can be Very Misleading [see Bauer & Koenig (2006)]
  - Consider that a Trial with "Trend"(e.g., p=0.09) may be Useful Supportive Evidence
  - With time-to-event primary endpoint consider possibility of Delayed Effect of Treatment
- Regulators generally do not allow α "buy-back" when a futility look is incorporated

Interim Analysis for Early Demonstration of Efficacy in Oncology

- Group-Sequential Designs (GSD) often beneficial in oncology and can Allow the Possibility of an Earlier Submission
- Generally preferable to Define Interim Analyses (IAs) in Terms of Target Numbers (TN) of Events
- Changing their Set Up (number, timing, spending fn.) can often Give Large Gains
  - See Smith (2006) for example where such changes reduced expected duration of trial by 3.2 months
- Looks After Equal Propns of events is often Not the Best Option, e.g., having IAs at 100, 200, and final at 300 may be problematic

Interim Analysis for Early Demonstration of Efficacy in Oncology, cont'd

- No Point in having 1st IA Before reach an Adequate Safety Database Size (for both S/T and L/T)
  - 1st Interim may therefore occur at X Events AND a particular Minimum trial (or follow-up) Duration
- If primary endpoint is PFS, may also want to build in OS requirements to stop at interim,
  - e.g., unlikely (low PP or CP) to obtain statistical significance on OS if trial continues to final analysis
- Best Times for Early Efficacy IAs generally not the same as Best Times for Futility IAs.
  - Some Adaptive Design approaches force these looks to occur at the same time

Adaptive Re-Sizing Based on Interim Effect Size

- Blinded Look(s) with Time to Event Endpoints
  - Power is guaranteed (under prop. haz.) at TN of Events
  - Blinded interim look is still useful, to estimate when EOS will occur, and then extend/ramp-up enrollment if indicated
- Adaptive Re-Sizing Based on an Interim Effect Size
  - Can be used to increase the TN of events if significance is projected to be borderline at EOS
  - This can be achieved by increasing # of pts, increasing TN of events (w/o increasing # of pts), or both
  - For simplicity I will just cover Sample Size Re-Estimation (SSR) based on the interim effect size, with just 2 stages

Adaptive Re-Sizing Based on Interim Effect Size, cont'd

- Define overall z-statistic (Z) by:

$$Z = [w]^{0.5} Z_1 + [1\text{-}w]^{0.5} Z_2$$

where Z_i is z-statistic from stage i = 1, 2

let n_i = # of pts per group in stage i planned at outset

& n_{2R} = revised # of stage 2 pts (a function of Z_1)

- The many variants of SSR basically fall under 2 main categories:
  - Weight all n_1+n_{2R} pts equally (and revise α to control T1E), i.e., Base on a Sufficient Statistic
  - Use original weights ($n_1/[n_1+n_2]$), and still test at original α level

Adaptive Re-Sizing Based on Interim Effect Size, cont'd

- Regulators require full pre-specification of algorithm for determining n_{2R} (and method for T1E control)
- Potential problem then as sponsor could back-calculate interim effect size
  - Could lead to different type of patients before/after adaptation
  - Impact reduced if investigators do not know algorithm
  - Problem can be reduced if new sample size can only take 2 or 3 different values [Gallo et al, 2006]
  - May need to show consistency of effect size before and after adaptation for some regulatory agencies

Adaptive Re-Sizing Based on Interim Effect Size, cont'd

- Some open issues related to use of Adaptive Re-sizing in the Oncology Setting:
  - When to increase TN # of events via increase in # of pts only, by increase in trial duration only, or by a combination
    - Would like to avoid need for suspension of enrollment if possible
  - Optimal timing of "look" at end of stage 1
  - Extensive power comparisons of two main approaches
  - Real life examples showing gains (in time and/or power) from incorporation of SSR based on interim effect size
  - Which group (DMC or unblinded statistician) should implement this deterministic SSR algorithm

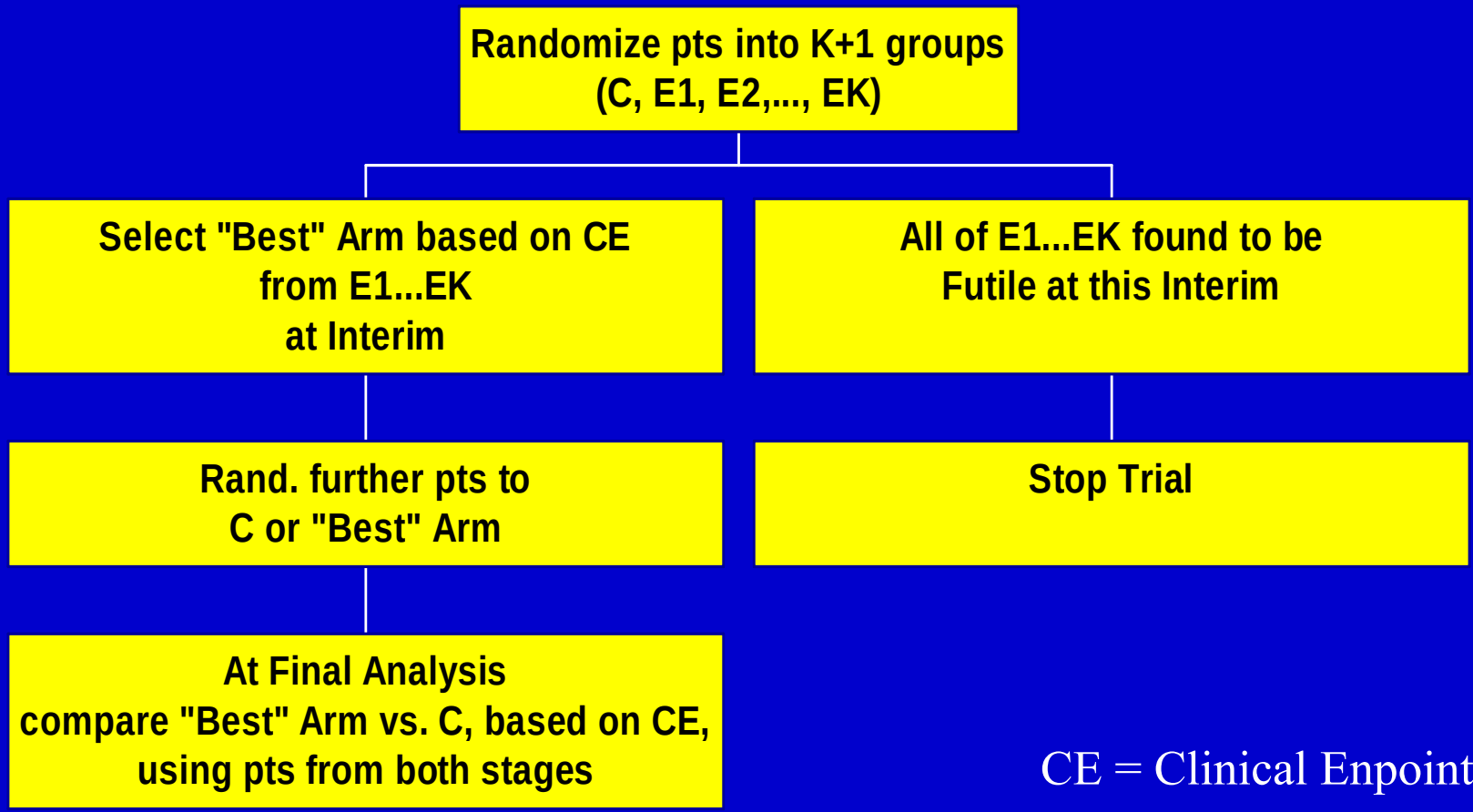
Determining the "Best" Dosage Regimen

- For cytotoxics MTD is established in Phase I
- In many other cases the Experimental Arm(s) to Take Into a Phase III Oncology Trial are Uncertain. We may need to decide on:
  - Different Doses of a Cytostatic or
    Different Doses/Schedules for a Cancer Vaccine
  - Different Dosing Regimens of a Cytotoxic
  - Different Combination Regimens
- MTD dose escalation approach is not useful in many such cases

Determining the "Best" Dosage Regimen, cont'd

- Traditional Approaches tended to Study Each Arm Separately in a Single Arm Trial based on Response Rate [A]
- Other Possible Approaches:
  - Randomized Parallel group Ph II design based on RR or Propn w' Progression by Set Time [B]
  - Skip Ph II & Take 2 or more Dosage Regimens into Ph III [C]
  - Skip Ph II, "Guessing" Single Dosage Regimen to use in Ph III [D]
- A, C, D Contribute to the High Oncology Attrition Rate
- B may be Preferable, but Lengthens Drug Devel. Time
- Many companies instead considering Phase IIB/III Combination Designs

Phase IIB/III Combination Designs Incorporating Treatment Selection & Testing



CE = Clinical Enpoint

Thall et al (1988), Schaid(1990), Smith (2002, 2003), Stallard & Todd (2003)

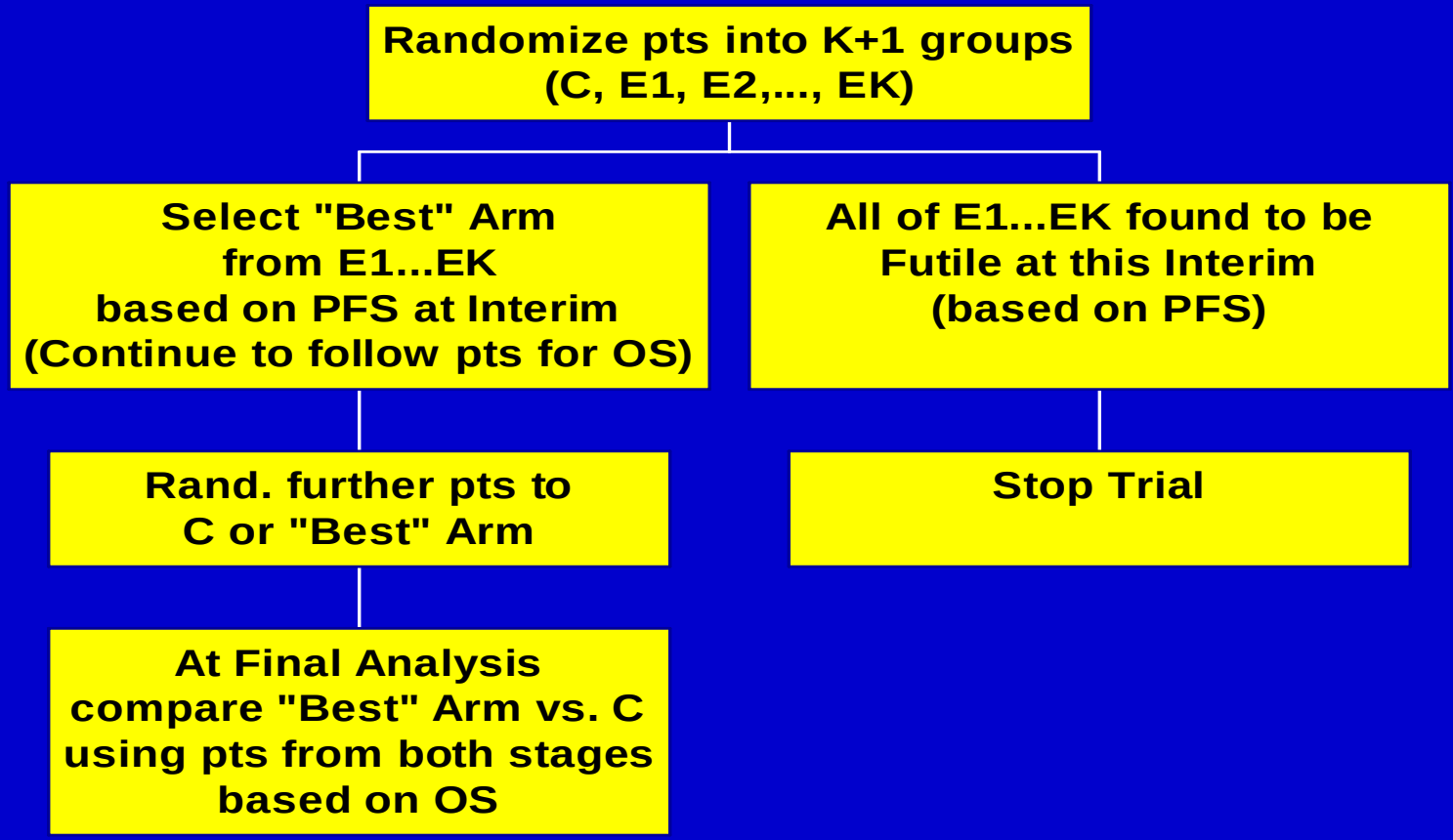
Phase IIB/III Combination Designs Incorporating Treatment Selection & Testing, cont'd

- Phase IIB/III Combination Designs particularly useful in oncology where often only 1 Ph III would be needed (maybe w' $\alpha = 0.010$)
  - If so we must have supportive evidence for submission from outside of this trial
- This Approach has Selection Carried Out on Clinical Endpoint (CE):
  - Requires CE to have a Short Median Time (relative to duration of enrollment), e.g.,
    - OS with metastatic RCC - medians 6-8 mths

Phase IIB/III Combination Designs Incorporating Treatment Selection & Testing, cont'd

- With Less Rapidly Progressive Cancers, we Can Still Use Ph IIB/III Design if there is an Intermediate Endpoint (IE) that:
  - IE occurs Early Enough AND is Highly Correlated with the CE
  - Stone (2007) and Royston (2003) each describe examples with IE = PFS and CE = OS
    - Royston's 1st line advanced ovarian cancer example has PFS medians 10m-13m, and OS medians 24m-29m
- For other cases (cancer types, stage, drug class) IE could be based on:
  - RR or propn. w' PD by a set time (provided it is highly correlated with CE)
  - PSA velocity (early predictor of PFS in Prostate Cancer)

Phase IIB/III Combination Designs Incorporating Treatment Selection & Testing, cont'd



see Royston et al (2003), Smith (2004), Todd & Stallard (2005), Stone (2007)

Phase IIB/III Combination Designs: Some Practical Considerations

- Avoid Using Ph IIB/III as an Excuse For Less Clear Objectives
- Important to Have Some Information Prior to Ph IIB/III on How Each Arm Performs in this pt population
- May want to Incorporate Rules under which Trial would Stop at End of Ph IIB (w' re-design), rather than continue into Ph III stage
- Timing of Interim has to Consider and Balance:
  - High Power needed in Current Study
  - High Chance that "Good" Dosage Regimen is Selected
  - Time of IA is late enough for Sufficiently Good Dose-Response Information to be obtained (if applic.)
  - As Few Pts as possible Given Sub-Optimal Dosage Regimen

[See Smith (2003) for further details]

Phase IIB/III Combination Designs: Some Regulatory Considerations

- Request very Early Discussion with FDA/EMEA if Planning to use a Ph IIB/III Combination Design
  - Discuss at End of Ph IIA Meeting w' FDA
  - Determine if supportive information is sufficient?
  - Obtain buy-in to any IE that is being used for selection
  - Agree on method for T1E control and on required overall α level in Ph IIB/III

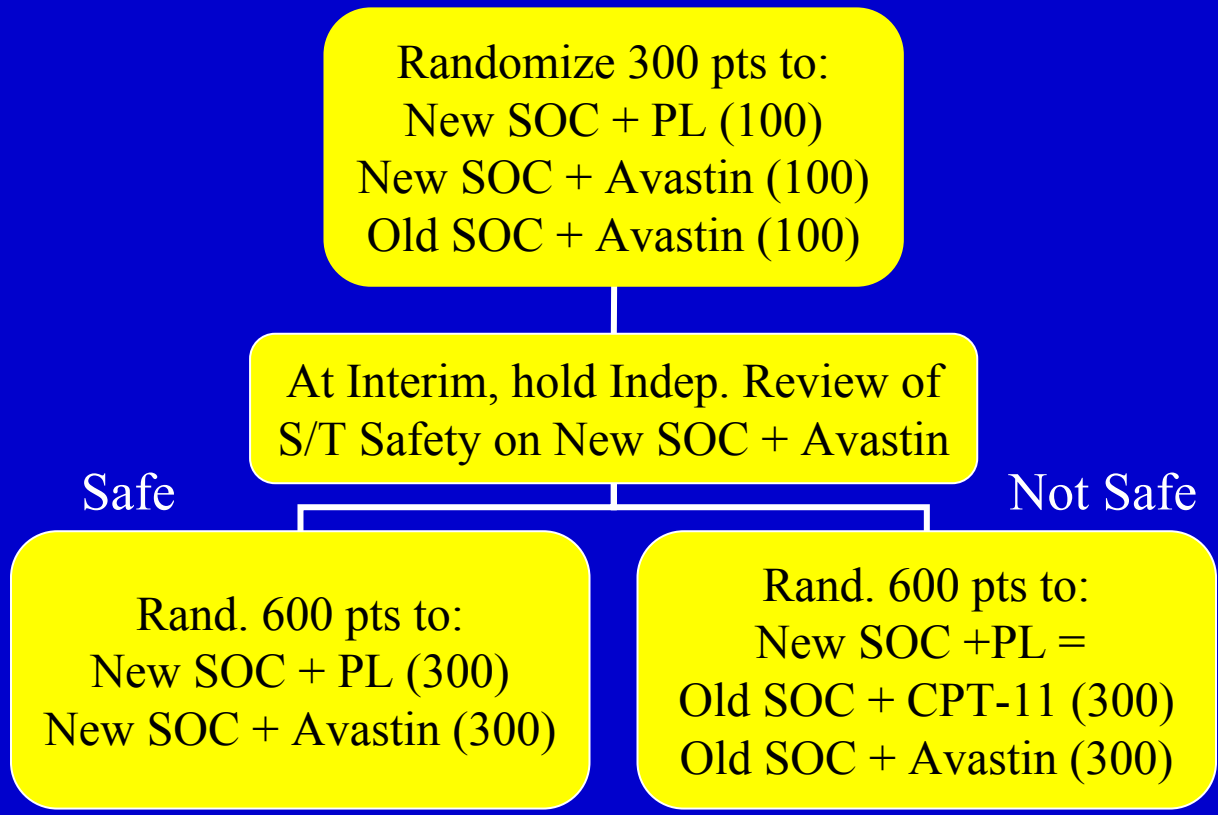
Phase III Designs with Safety-Based Treatment Selection

- Temple (2004) has argued that for some Phase III Outcome-Based Trials it is best to Start with 2 Doses & Select Best Based Solely on Safety at an Interim
  - He often cites anti-platelet example where after interim:
    - Drop LD if HD has acceptable bleeding rate
    - Otherwise drop HD
  - Not useful in oncology for cytotoxics if MTD already established in Ph I
  - Potentially useful for trials of cytostatic/cytotoxic combination therapy
    - Few examples seen in oncology except for a related approach used on an Avastin trial

Phase III Designs with Safety-Based Treatment Selection, cont'd

- Hurwitz (2004), Bajamonde (2007) describe an example with a cytotoxic/cytostatic trial in metastatic CRC:
  - Safety and Efficacy of Avastin shown in Phase II as add on to 5-FU/LV (Old SOC)
  - Since then 5-FU/LV/**CPT-11** has been Approved and become the New SOC
  - No Safety Data for Avastin as Add-On to New SOC
  - To avoid waiting on a new Ph II assessing Avastin + New SOC, sponsor embedded Phase II-type decision in a 3-arm trial

Phase III Designs with Safety-Based Treatment Selection, cont'd

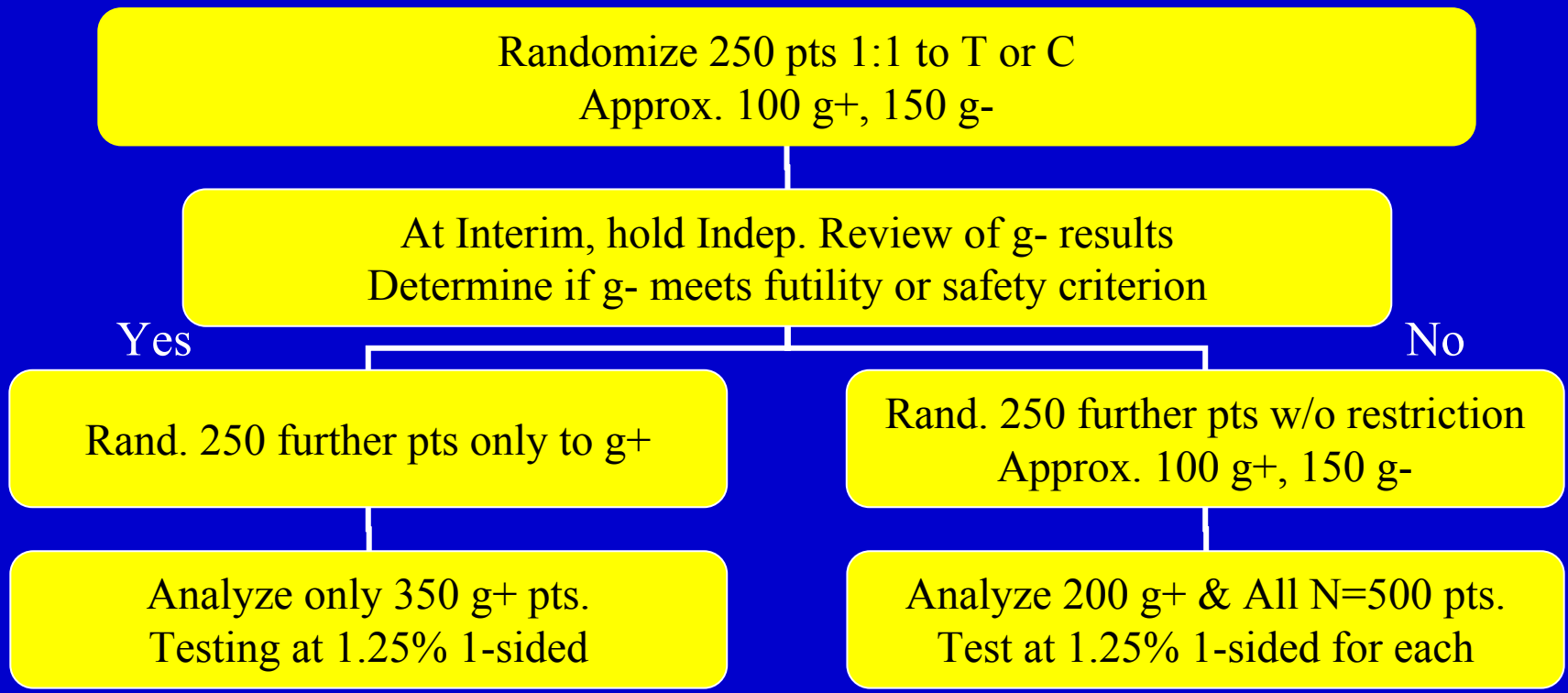


Final Analysis based on OS using all 400 pts per Arm.
Minimal Penalty for this type of Selection

Adaptive Selection of Patient Population

- In oncology it is well known that there are many compounds which only work in a subgroup of patients (e.g., those who express the known target), and many more where the effect is much greater in a subgroup
  - Subgroup can sometimes be determined by a Single Biomarker based on drug's (presumed) MOA, e.g., HER2-neu expression for herceptin
  - Subgroup sometimes based on a pharmacogenomic profile
- Adaptive design is sometimes useful here, including approaches which:
  - Select from 2 Subpopulations (e.g., hormone resistant or hormone refractory) at interim and Only Continue with "Best" in Stage 2
  - Determine at interim whether to continue with Whole Population or with just a pre-defined Subgroup in Stage 2
  - Adaptive Signature designs which develop a genomic classifier in Stage 1

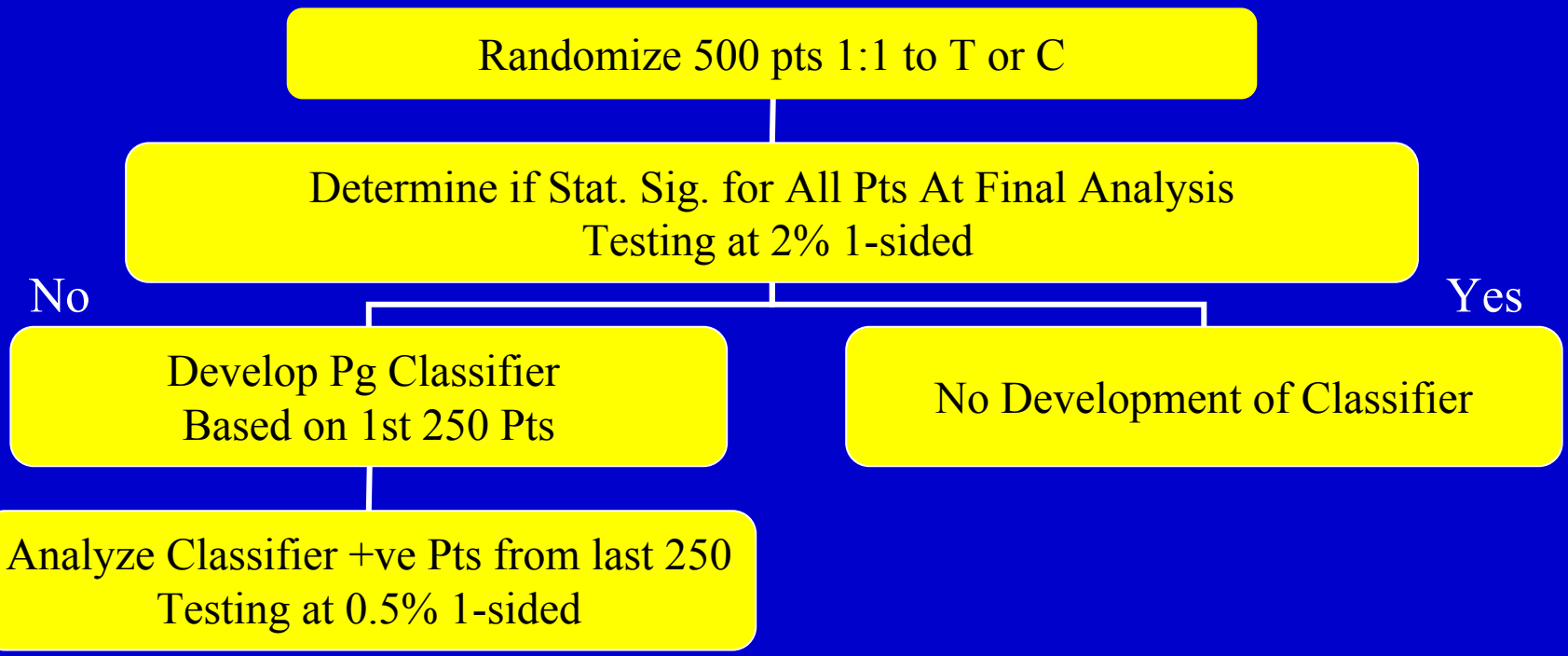
Adaptive Selection of Patient Population on Pharmacogenomic Basis: Wang et al (2007) Approach



Key point here is that Subgroup, α-Splitting, and Adaptation are All Fully Pre-Specified.

Above is just one example. Variants would use different N, α splitting, and interim fractions, depending on prevalence, ratios of assumed effect sizes, etc. Other variants are to incorporate Hochberg. Also applicability depends on time to read on efficacy, enrollment time for All, and enrollment time for g+.

Adaptive Signature Designs - Freidlin & Simon (2005), Simon (2006)



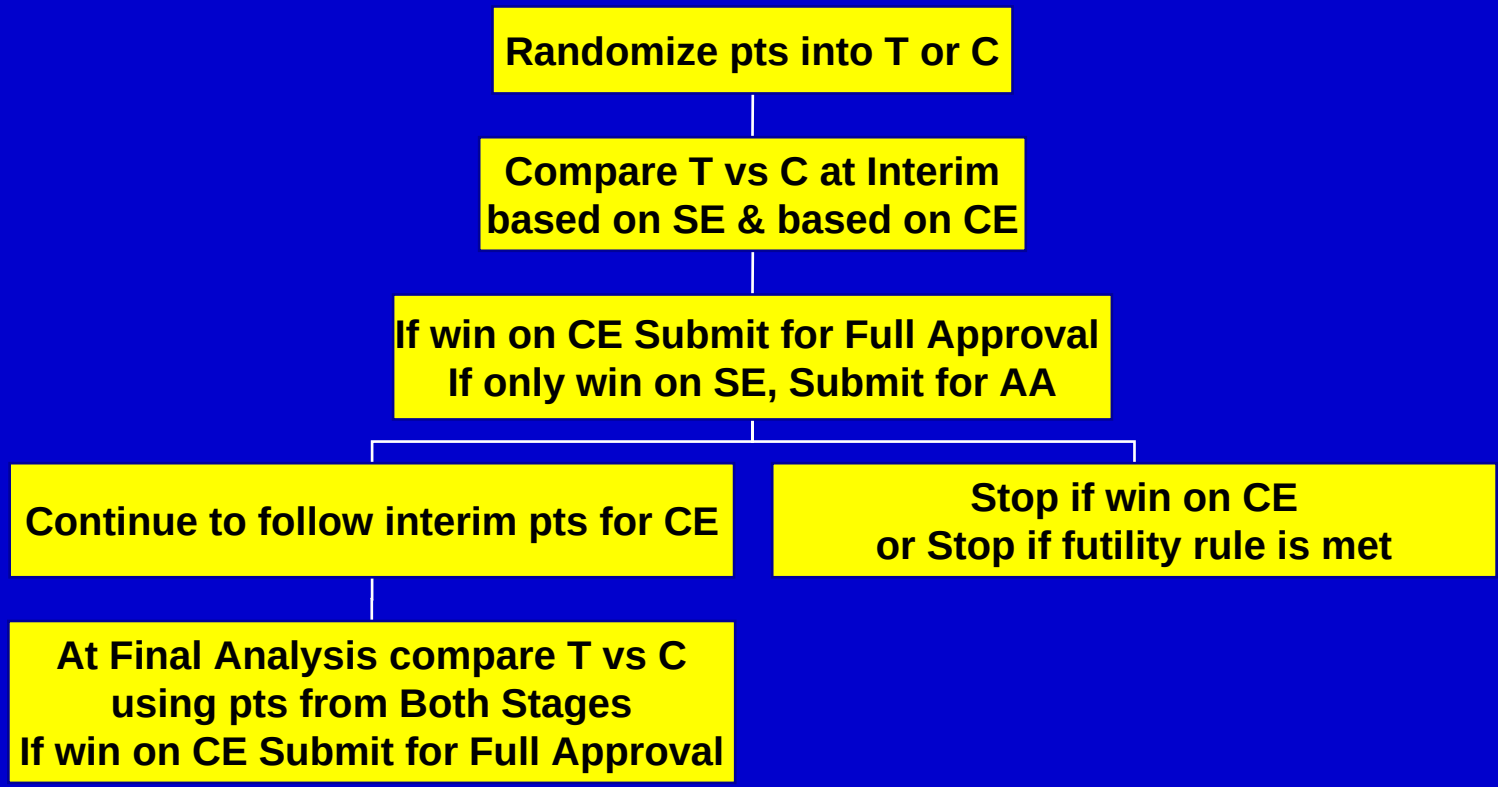
α-Splitting and Interim Fraction are Fully Pre-Specified,
but Classifier is Not Pre-Specified.
Acceptable to Regulators in Phase II, but May Not Be Acceptable in Phase III.

Variant of above method always develops classifier - runs risk of narrower label.

Phase II/III Designs for Both Accelerated Approval & Full Approval

- Accelerated Approval (AA), based on a Surrogate Endpoint, in oncology has historically been based mostly on RR from 1 arm trial(s)
  - Later Comparative Trial Required to Demonstrate Clinical Benefit, but problematic to complete such trials
  - Only 13/29 AAs (as of 2005) had completed trials to assess this clinical benefit [Ross, 2005]
- In last few years CDER has advocated the following approach:
  - **Use Single Randomized Trial for both AA and Full Approval**
  - AA based on IA of Surrogate Endpoint (SE)
  - Full Approval based on Final Analysis using CE
  - This Design Approach used first for Oxaliplatin in 2nd line therapy for metastatic colorectal cancer

Phase II/III Designs for Both Accelerated Approval & Full Approval, cont'd



CE = Clinical Endpoint, SE = Surrogate Endpoint, AA = Accelerated Approval
C = Control, T = Test

Phase II/III Designs for Both Accelerated Approval & Full Approval, cont'd

- Important to obtain prior FDA buy-in on proposed α levels to be used for SE and for CE
- Many variations on this theme including:
  - Whether or not to also assess CE itself at IA (may not have sufficient safety database for full approval at this time)
  - Whether/When to include extra interims to assess CE + how to spend α
  - Nature of futility rule and whether based on SE or CE
    - Whether to force trial to stop if not win on SE
  - Strategies re Enrollment (when CE needs more pts than SE)
    - Enroll aggressively until # of pts needed for CE
    - Enroll aggressively until # of pts needed for SE (w' enrollment suspension)
    - Enrollment (for SE pt #) timed to complete at IA - allows extension after interim

Stated Regulatory Requirements With All Types of Adaptive Design in Phase III

- Fully Pre-Specify Adaptive Methods (Designed Flexibility)
  - Limit amount of Adaptation in any Pivotal Trial
- Show that Type 1 Error is Controlled
- Provide Bias-Adjusted Estimates and Bias-Adjusted CIs
- Have Adaptation Performed by an Independent 3rd Party
- Ensure Implementation Avoids Operational Bias and Restricts Sponsor Exposure to Unblinded Results

O'Neill (2006), Wang (2006, 2007), FDA's DSMB guidelines (2005), EMEA's Flexible Design guidelines (2006)

Other Important Considerations With All Types of Adaptive Design in Phase III

- Discuss with Regulatory Authorities at an early stage if any Adaptive Design is being considered
- Useful to have FDA's Special Protocol Assessment
- Important to get full DMC (or other group carrying out adaptation) buy-in to any decision rules by which they will be guided
- Important to demonstrate gain vs non-adaptive or simpler GSD approach

Concluding Comments on Use of Adaptive Design in Pivotal Oncology Trials

- Oncology is an area in which Adaptive Design has Already been Used Extensively for several years
- Seamless Phase IIB/III, Adaptive Re-Sizing based on interim effect size, Adaptive Selection of Patient Population (particularly on a pharmacogenomic basis), and Phase II/III for AA/Full Approval, can each be advantageous in oncology
- Important to compare vs simpler approaches such as GSDs and to ensure that gains outweigh the extra complexity
- Adaptive Design approaches can, in many cases, help **Cut Attrition Rate** and/or **Cut Costs** of late stage Oncology Drug Development