# Further Developments in the Use of Combination Phase II/III Designs to Accelerate Drug Development

- Jonathan R. Smith, Ph.D.

June 18th, 2003, DIA Annual Meeting, San Antonio

# Outline of Presentation

- Potential uses of Phase II/III combination designs
- Derivation & Properties of 2 look Phase II/III designs
  - Type I error control
  - Power and Sample Size calculation
  - Optimal Timing of Interim
- Comparison vs. separate Phase II & Phase III
  - Sample Size savings
  - Time Savings
- Future Work
- Conclusions

# Situations Where Adaptive Designs are Most Useful

To obtain maximal benefit from use of interim analysis or adaptive design we would like:

1. Length of Follow-up to be short relative to duration of enrollment
2. Short time from LPV to Interim Analysis decision
   - EDC can help cut this time
3. Randomization via IVR if arms can be dropped

Note: If follow-up is long but onset of action is rapid we can overcome #1 by using results on the primary endpoint from an earlier visit, or by using a surrogate.

# Phase II/III Combination Trial - Situations Where this could be Useful

- Range of doses will likely cover optimal dose
  - so that arms will not need to be added
- At most 6 doses still under consideration
- Major safety concerns not likely to apply to very many doses
  - otherwise separate Phase II is probably preferable

# Phase II/III Combination Trial - Situations Where this could be Useful, cont'd

- Certain Single Study Submissions with a dose-ranging component (provided that SSS has sufficient outside "confirmatory evidence")
  - Large multicenter study (with the same Phase II, III endpoints)
  - New stage of disease, closely related disease
  - New patient population
  - New combination therapy
  - Orphan indication or other rare disease

# Phase II/III Combination Trial - Situations Where this could be Useful, cont'd

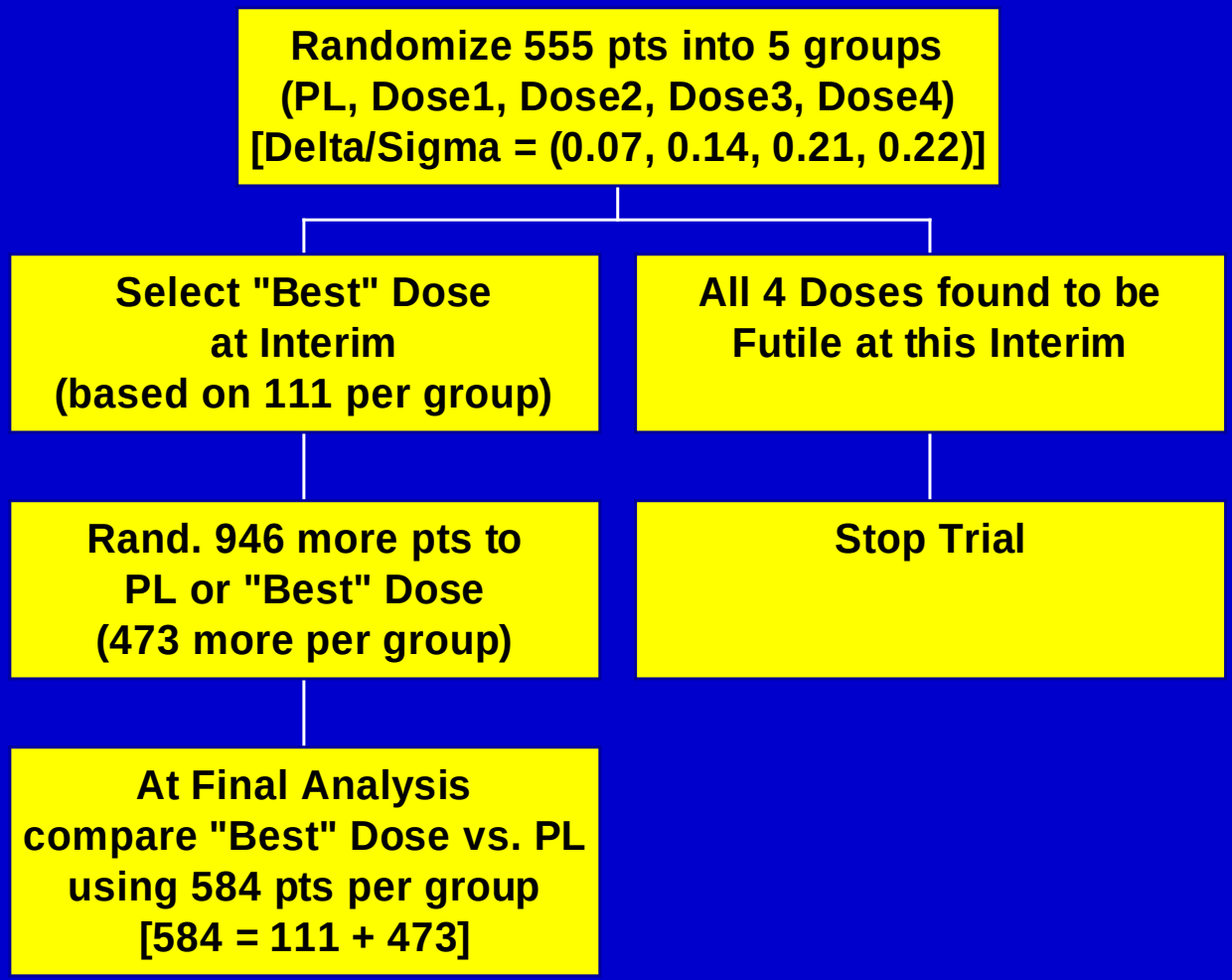
- Where sponsor would otherwise carry out Phase II, Phase III #1, Phase III #2 in sequence, perhaps due to limited funds
- In place of a multi-armed Phase III trial
  - e.g., in place of past "Phase III trial" that included 5 doses and Placebo

# Aims of Phase II/III Combination Studies Considered Here:

- Combine dose selection and confirmatory stages
  - start with 2-5 doses + placebo
- Select "Best" dose at Interim and continue with "Best" dose & Placebo
- Stop study early if all doses are determined at interim to be ineffective

Note: Could also incorporate additional interims to stop for success, as in Smith (2002), and Stallard & Todd (2003)

# Flowchart for Phase II/III Design



# Phase II/III Combination Design

1. At start randomize to PL, $D_1,\ldots D_k$
2. At Interim
   - Choose "Best" dose $D_{BD}$
   - *Decision rule specified in protocol & administered by independent group*
   - Stop if all $D_j$ futile
   - Randomize to PL, $D_{BD}$ from now onwards
3. At Final
   - Test $D_{BD}$ vs. PL at level $\alpha_2$
   - Efficacy demonstrated if test statistic $\geq Z_{1-\alpha 2}$

# Notation and Assumptions

- Assume data is normally distributed
  - Can also be applied to binary or survival data
- $Z_{iDj}$ = test statistic at look #i for $D_j$ vs. PL
- We test at 1-sided $\alpha_2$ at EOS
- $Z_{1BD} < Z_{1-\alpha 0}$ is futility decision rule used to stop study at interim (with corresp. CP)
- $n_1$ per group at interim; $n_2$ per group by EOS; with $\tau = n_1/n_2$ the information fraction

# Experimentwise Type I Error

Type I error is given by

$$\sum_{j=1}^{k} P( Z_{1Dj} \geq Z_{1-\alpha_0} \cap Z_{2Dj} \geq Z_{1-\alpha_2} \cap D_j \text{ selected} )$$

Note: this applies whatever decision rule is used to select $D_j$

# Experimentwise Type I Error, cont'd

Suppose decision rule at interim is to choose $D_j$ corresponding to highest $Z_{1Dj}$ .

$$\text{T1E} = \sum_{j=1}^{k} P(Z_{2j} \geq Z_{1-\alpha 2} / Z_{1j} = \max(Z_{11},\ldots, Z_{1k}))$$

$$* P(Z_{1j} = \max(Z_{11},\ldots, Z_{1k}))$$

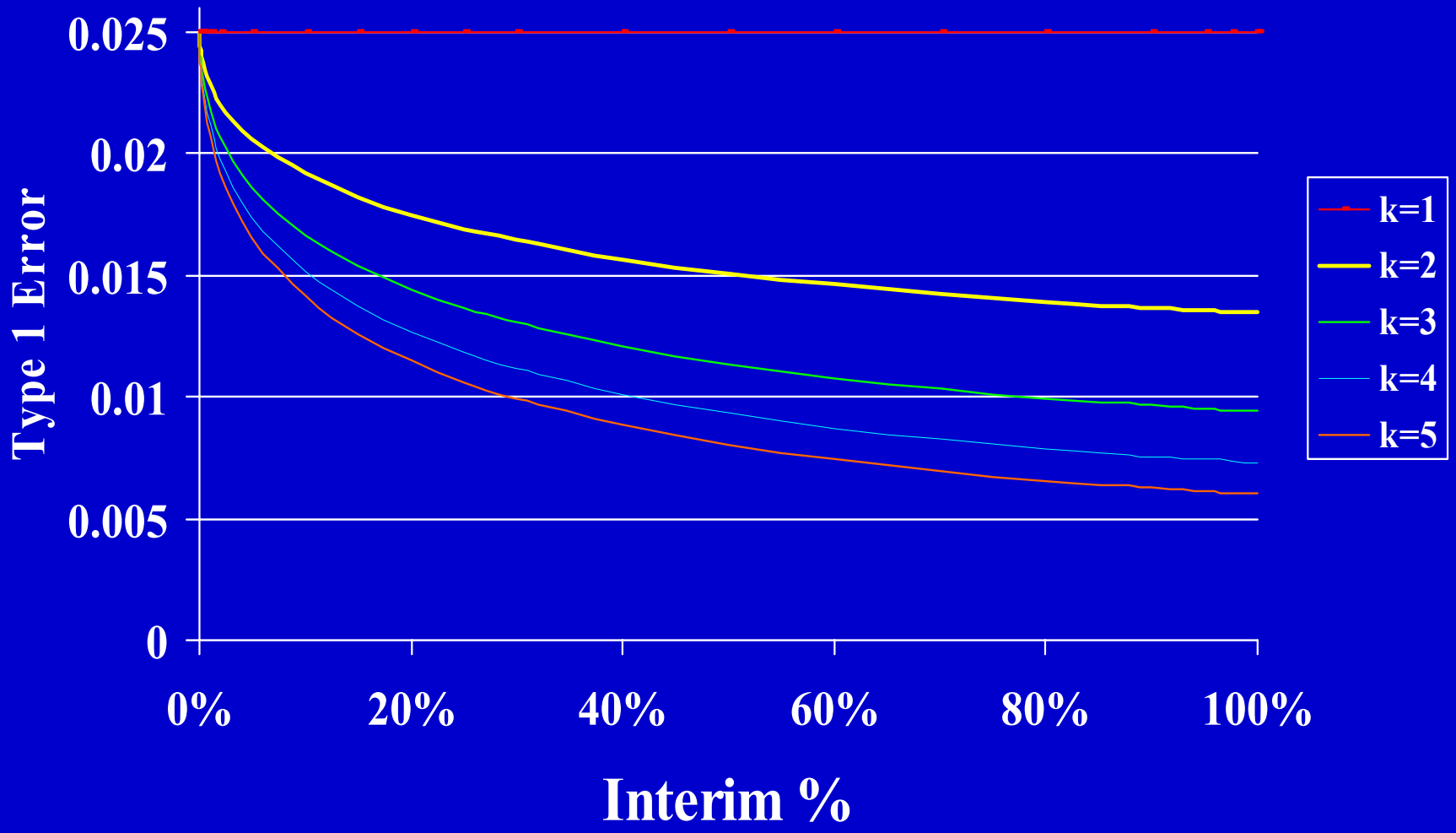
Note: In case sponsor overrides DSMB recommendation to stop trial for futility, $\alpha_2$ calculation here has not made use of built in futility look to increase $\alpha_2$ further.

See also Simon et al (1994), Hsu et al (1997), Smith (2002), Stallard & Todd (2003), Soo, Lan & Shun (2003)

# Calculation of Critical Alpha levels

- Equate T1E expression to $\alpha = 0.025$ (1-sided)
- With equal replication of active arms this can be expressed as a 1d integral (at least 2 different ways)
- Evaluate 1d integral numerically
  - Higher dimensional integrals needed when actives are not equally replicated
- For given $\tau$, k, $\alpha$, solve (iteratively) for $\alpha_2$

# Critical Alpha Value as a Function of Timing for Interim and # of Active Arms



# Critical Alpha levels

| k | 0.1% | 10% | 20% | 50% | 99.9% | Dunnett's alpha |
|---|---|---|---|---|---|---|
| 1 | 0.02500 | 0.02500 | 0.02500 | 0.02500 | 0.02500 | 0.02500 |
| 2 | 0.02428 | 0.01918 | 0.01751 | 0.01510 | 0.01350 | 0.01348 |
| 3 | 0.02393 | 0.01667 | 0.01443 | 0.01136 | 0.00944 | 0.00941 |
| 4 | 0.02370 | 0.01517 | 0.01266 | 0.00933 | 0.00733 | 0.00731 |
| 5 | 0.02353 | 0.01414 | 0.01147 | 0.00803 | 0.00603 | 0.00601 |

# Determination of Power & Sample Size

- Power
  - power is function of $\alpha_2$, k, $\tau$, $\Delta$, and $N_{tot}$
    - where $N_{tot} = n_1*(k+1) + 2*(n_2-n_1)$
  - evaluated analytically here via 2k, 1d integrals
- Sample Size (SS)
  - Optimal Phase II/III design (for given $\alpha_2$, k, $\Delta$, and power) takes $\tau = \tau_{opt}$
    - where $S(\tau_{opt}) = \min_{\tau} S(\tau)$

# Example of Optimal Phase II/III Design

- k=4, $\underline{\Delta}' = (0.07\sigma, 0.14\sigma, 0.21\sigma, 0.22\sigma)$
- Minimum total sample size to give 90% power is $N_{tot} = 1501$
  - $\tau_{opt} = 0.19$
    - all other values of $\tau$ give a greater sample size
  - $n_1 = 113$ per group at interim
  - $n_2 = 581$ per group at EOS
  - here $\alpha_2 = 0.01276$ to guarantee T1E = 0.025

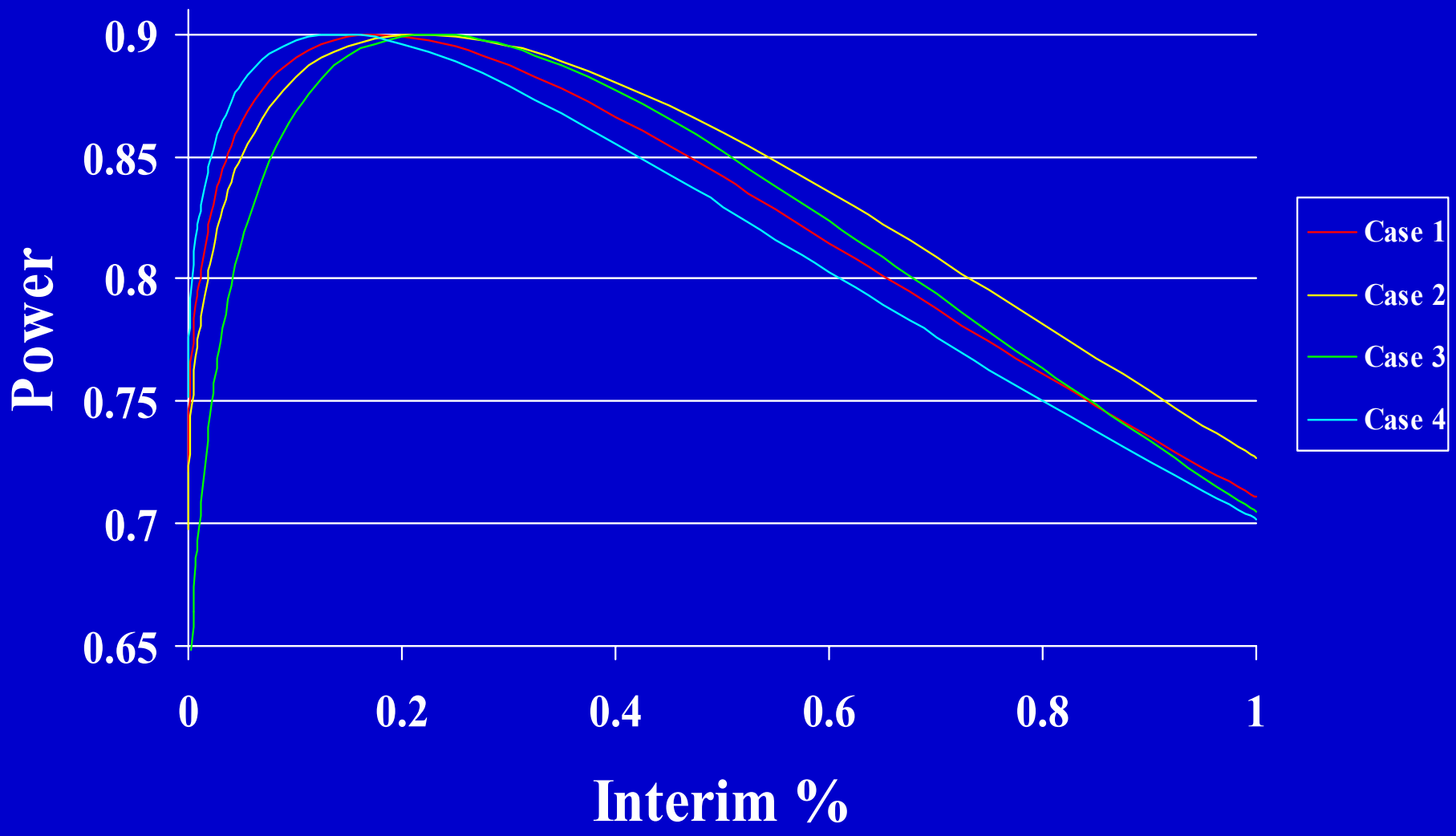
# Optimal Timing of Interim

- Investigated for k = 2, 3, 4, and 5 active arms
- k=4 active arms - 4 cases
  - $\underline{\Delta}' = (0.07\sigma, 0.14\sigma, 0.21\sigma, 0.22\sigma)$ asymptotic
  - $\underline{\Delta}' = (0.06\sigma, 0.12\sigma, 0.18\sigma, 0.24\sigma)$ linear
  - $\underline{\Delta}' = (0.03\sigma, 0.11\sigma, 0.20\sigma, 0.23\sigma)$ sigmoidal
  - $\underline{\Delta}' = (0.08\sigma, 0.16\sigma, 0.24\sigma, 0.16\sigma)$ curve turns

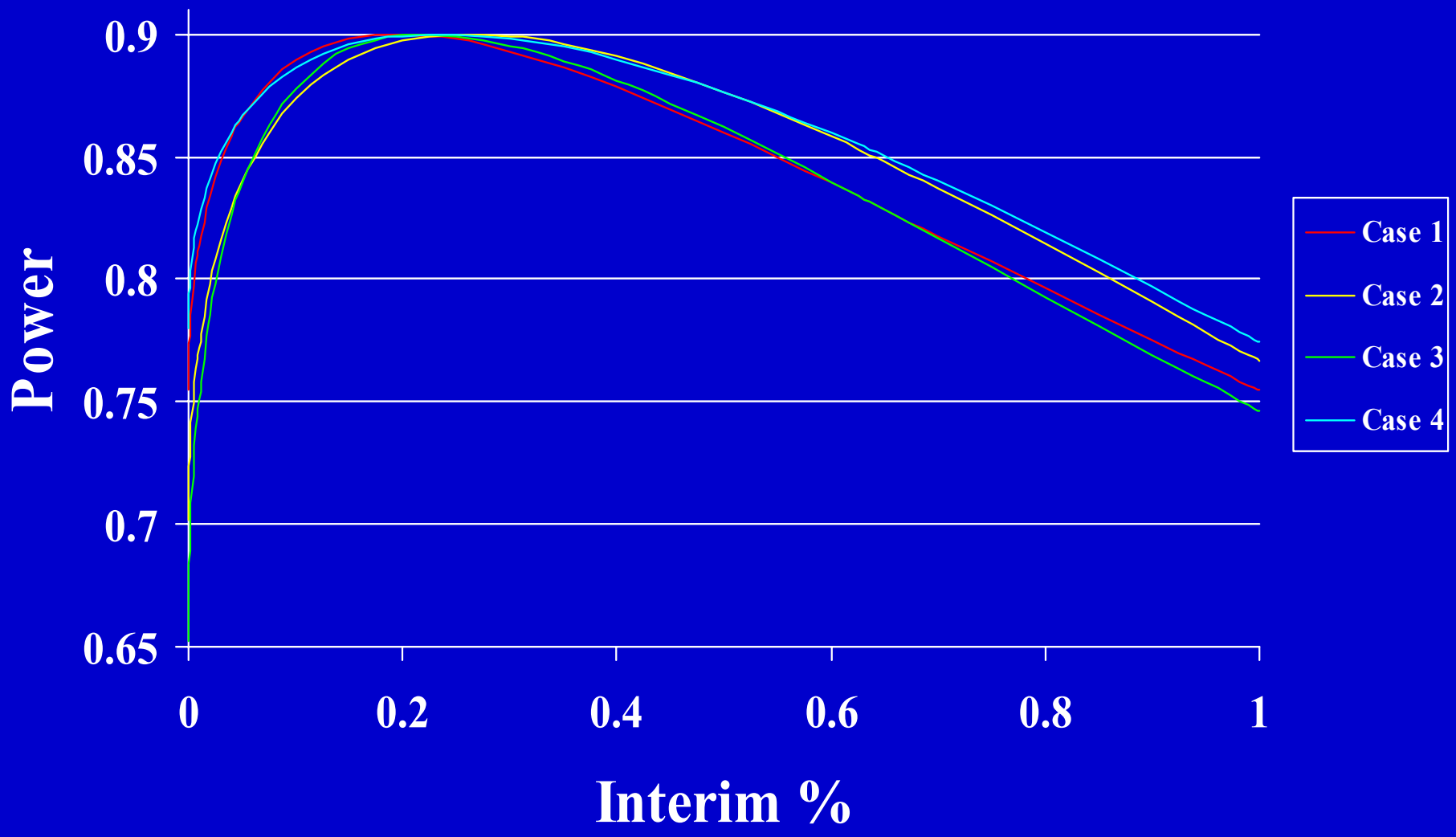
# Optimal Timing of Interim

- k=5 - four shapes similar to k=4
- k=3 - asymptotic, linear, curve turns
- k=2 active arms - 4 cases
  - $\underline{\Delta}' = (0.22\sigma, 0.22\sigma)$ $\quad \Delta_1/\Delta_2 = 1.0$
  - $\underline{\Delta}' = (0.185\sigma, 0.22\sigma)$ $\quad \Delta_1/\Delta_2 = 0.84$
  - $\underline{\Delta}' = (0.15\sigma, 0.22\sigma)$ $\quad \Delta_1/\Delta_2 = 0.68$
  - $\underline{\Delta}' = (0.11\sigma, 0.22\sigma)$ $\quad \Delta_1/\Delta_2 = 0.50$

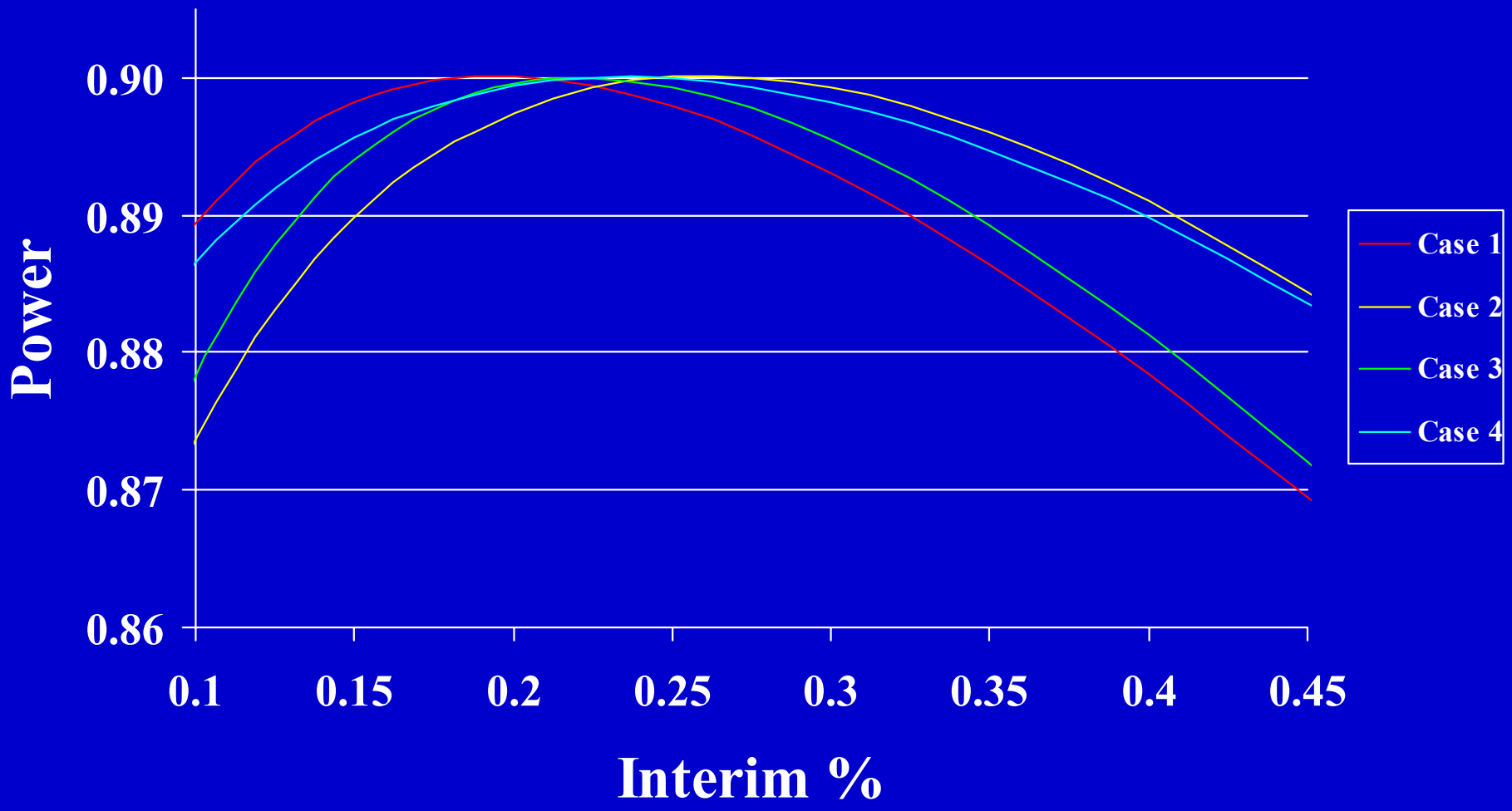
# Power as a Function of Timing of the Interim - K=5



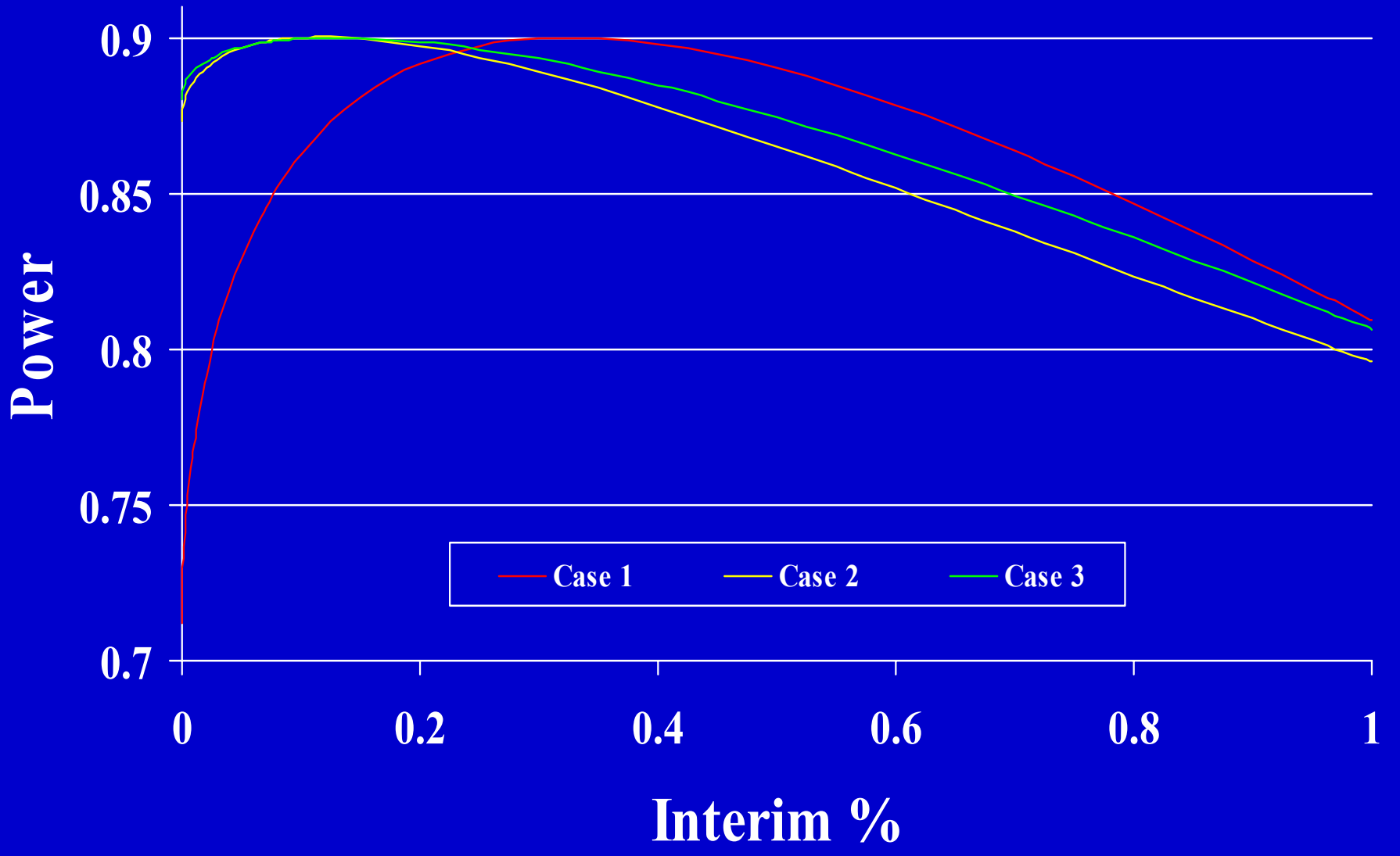
# Power as a Function of Timing of the Interim - K=4



# Power as a Function of Timing of the Interim - K=4



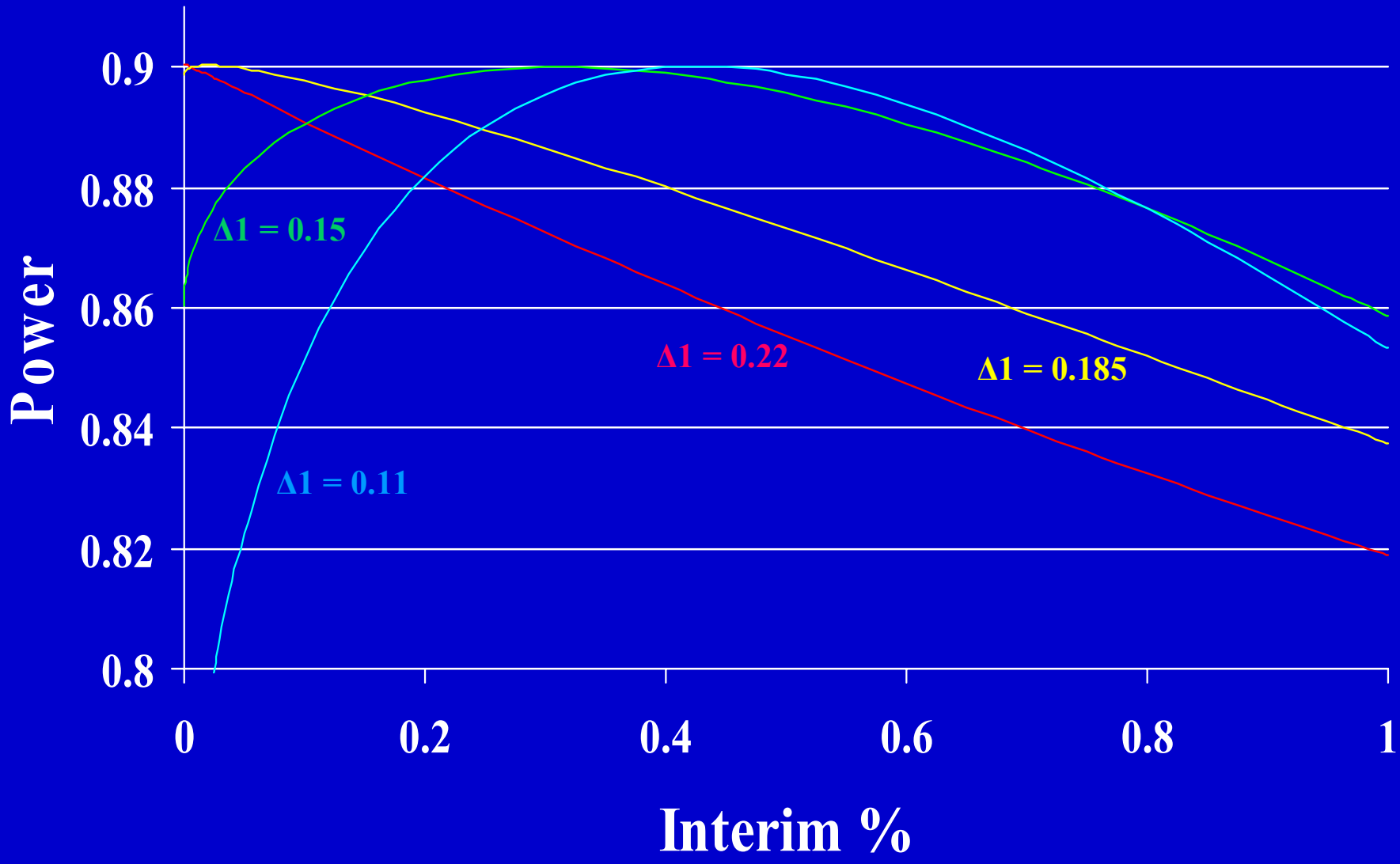
# Power as a Function of Timing of the Interim - K=3



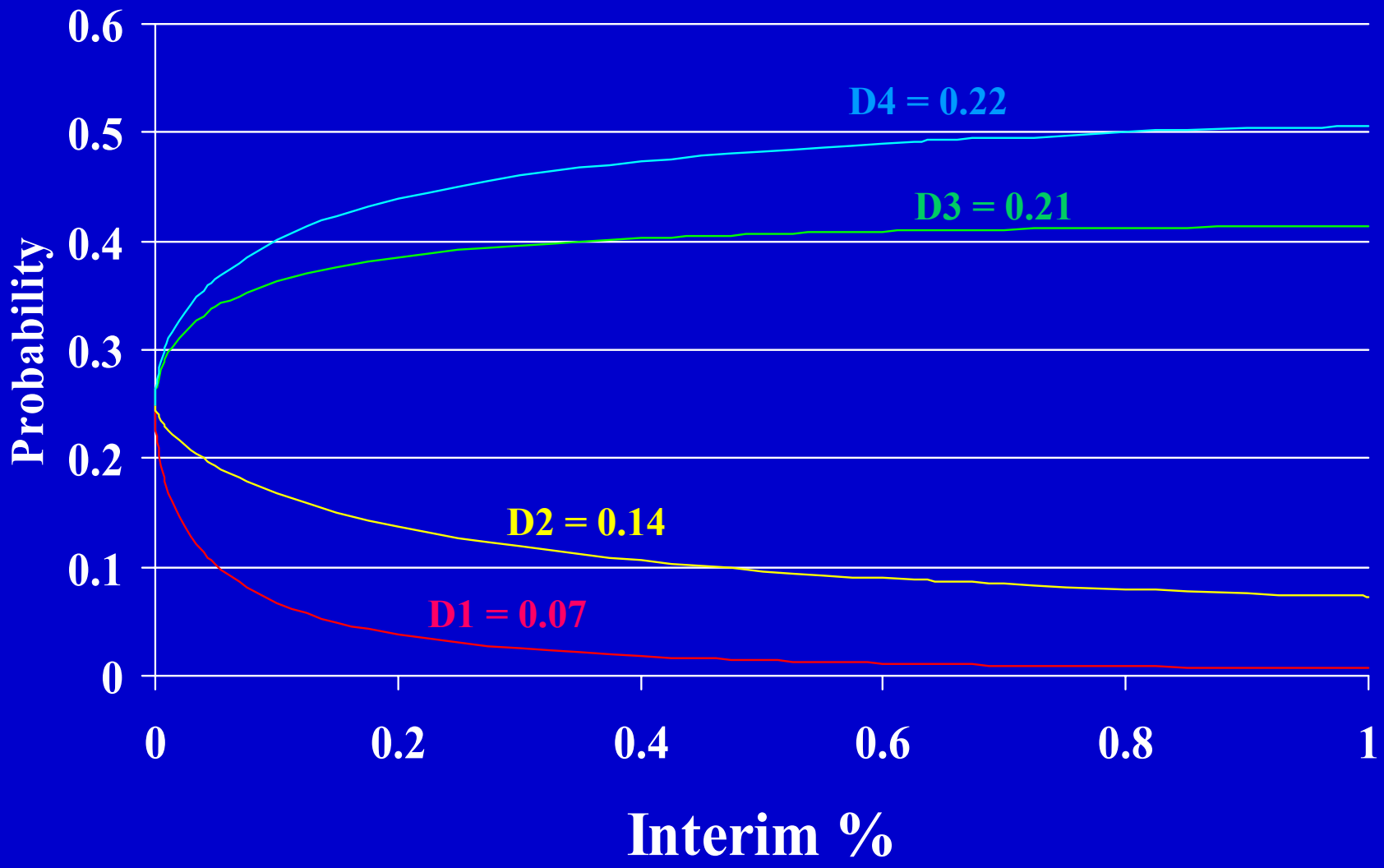
# Optimal Tau based on Power alone, k=5

| Case | $\tau_{opt}$ | Power $\tau = 0$ | Power $\tau = \tau_{opt}$ | Power $\tau = 1$ |
|---|---|---|---|---|
| 1 | 0.18 | 72.1% | 90.0% | 71.1% |
| 2 | 0.22 | 69.8% | 90.0% | 72.7% |
| 3 | 0.23 | 57.5% | 90.0% | 70.5% |
| 4 | 0.14 | 75.2% | 90.0% | 70.2% |

# Power as a Function of Timing of the Interim - K=2



# Probability of Selecting each Dose at Interim
## Case 1, K=4



# Timing of Interim

- When determining timing of Interim #1 seek to balance
  - (a) Need for high power in current study
  - (b) High chance that "best dose" is selected at interim #1
  - (c) Time of interim is late enough for good dose-response information to be obtained
- With k=4, $\tau_{opt} \in (0.19, 0.26)$, but $\tau$=0.32 - 0.41 (depending on curve shape) may be preferable when allow for (b)-(c)
  - retains power at 89%

# Timing of Interim

- When allow for (b) and (c) following would be preferable:
  - for k = 5, $\tau$ = 0.24 - 0.33 (depending on curve shape)
  - for k = 3, $\tau$ = 0.28 - 0.50 (depending on curve shape)
- For k = 2 (see also Soo, Lan, & Shun, 2003)
  - for $\Delta_1/\Delta_2$ in the range 0.5 - 0.68, having the interim at $\tau$= 0.6 still gives ≥ 89% power + keeping all 3 arms until EOS ($\tau$=1) gives ≥ 85% power
  - for $\Delta_1/\Delta_2$ in the range 0.84 - 1.0, $\tau$ = 0 gives best power, but no measure of dose-response

**Comparison of Phase II/III Designs with Separate Phase II and Phase III, based on**

**(a) Sample Size**

**(b) Timelines**

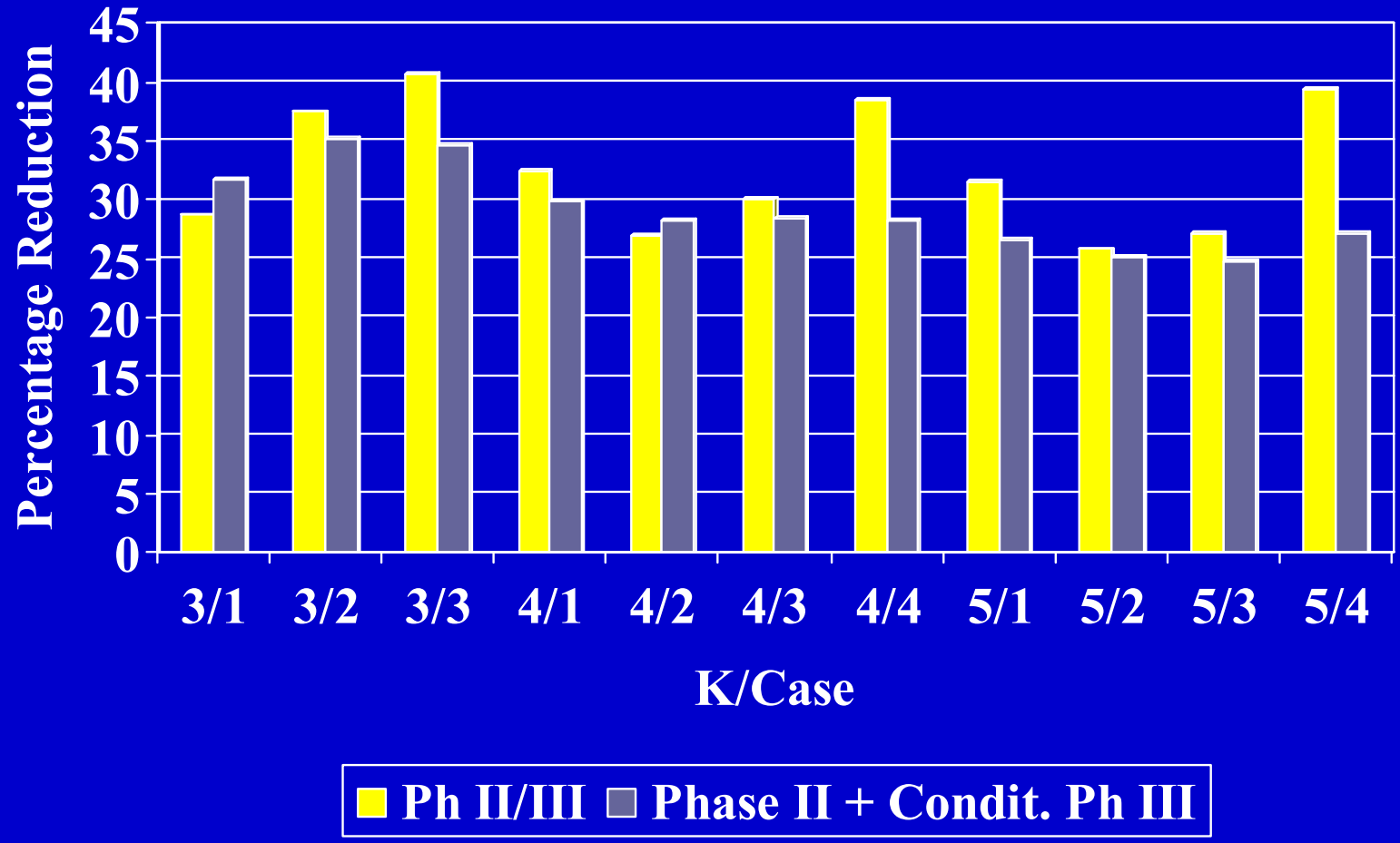
# Comparisons based on Sample Size

- K=3, 4, 5 with 3-4 cases each, as before
- In Phase II/III (with $0<\tau<1$) also have patients enrolled between last interim patient and implementation of dropping k-1 arms
  - Assume primary endpoint at 30 days
  - 4w from LPV to drop (k-1) arms
- Enrollment rate chosen to enroll Ph II/III in 12m
- Sample size for Phase IIs based on Williams' test with 80% power
- Assume 90d between Ph II's LPO and Ph III's FPI

# Comparisons based on Sample Size, cont'd

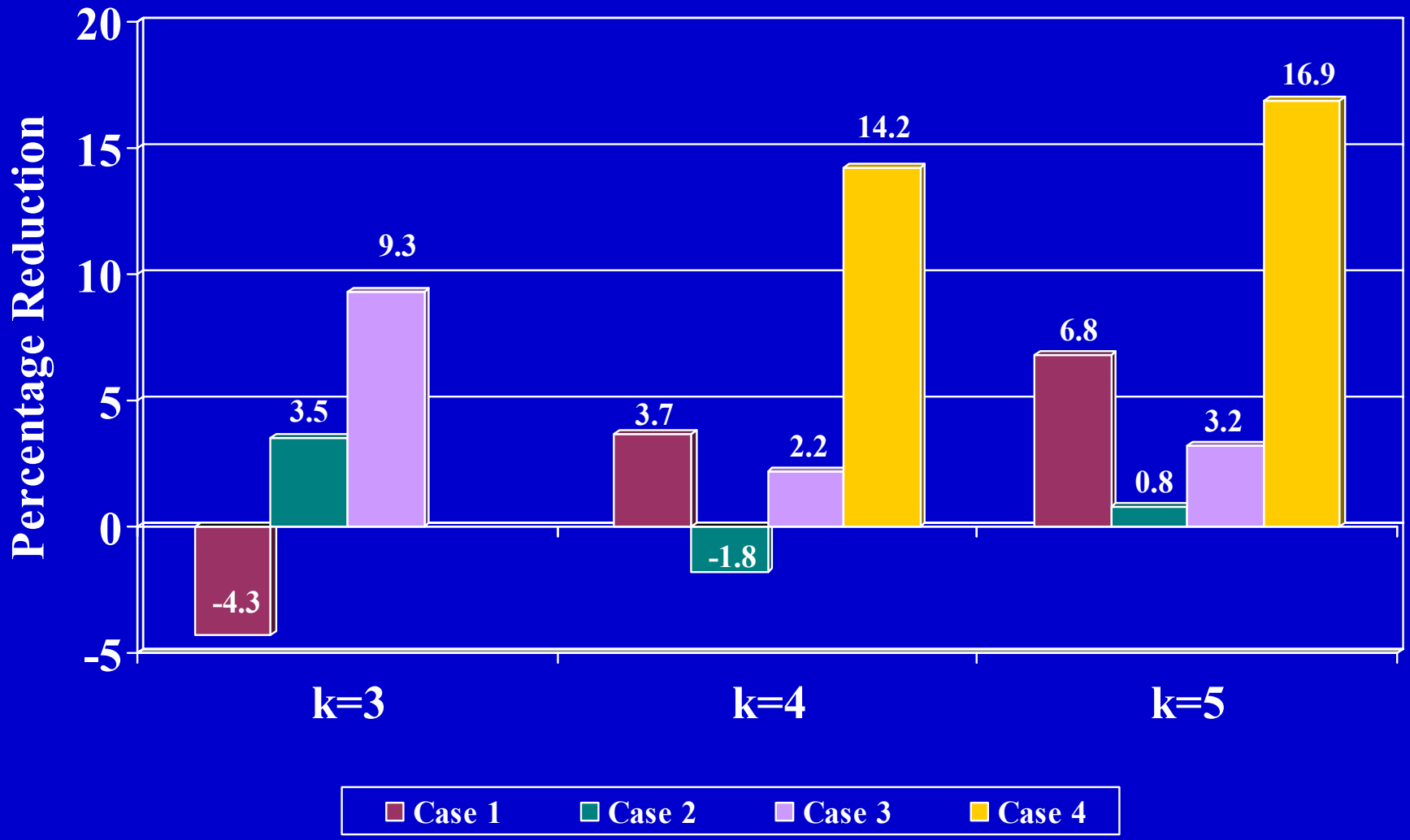
- Scenario 1
  - Phase II not carried out to Phase III standards - always followed by Phase III
  - Phase III based on arm with max Z from Phase II
- Scenario 2
  - Phase II is carried out to Phase III standards - only followed by Phase III if all primary comparisons are non-significant (using Step-down Dunnett's for MC)
  - Phase III based on arm with max Z from Phase II, but using max amongst those with non-significance

# Percentage Reduction in Sample Size vs Scenario 1 (Phase II + Phase III always)



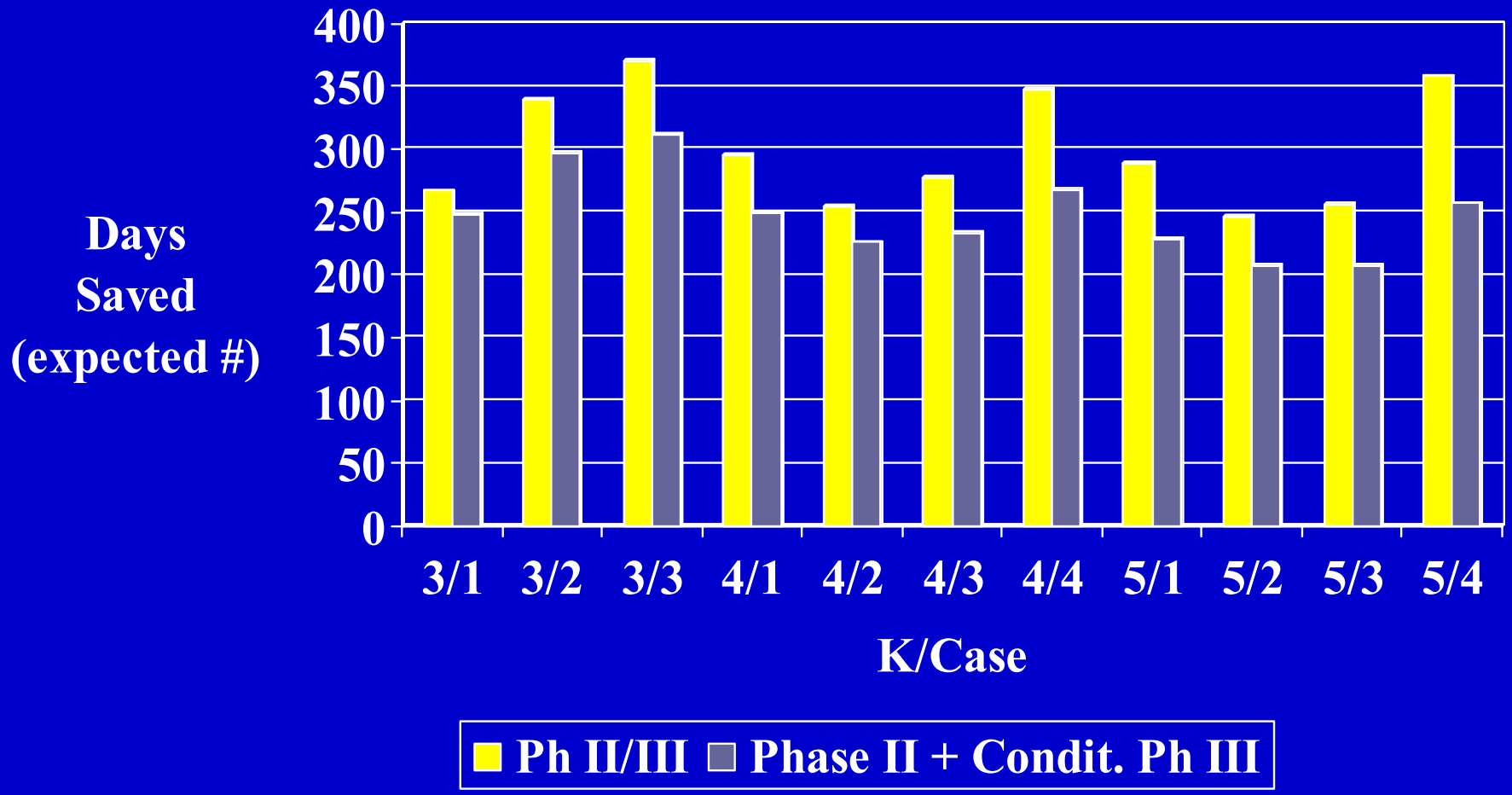

Note: Ph II/III also gives large % reduction in # of pts vs Scenario 1 in all 11 cases with 60d, or 90d primary endpt.

# Percentage Reduction in Sample Size from Phase II/III vs Scenario 2 (Phase II + conditional Phase III)



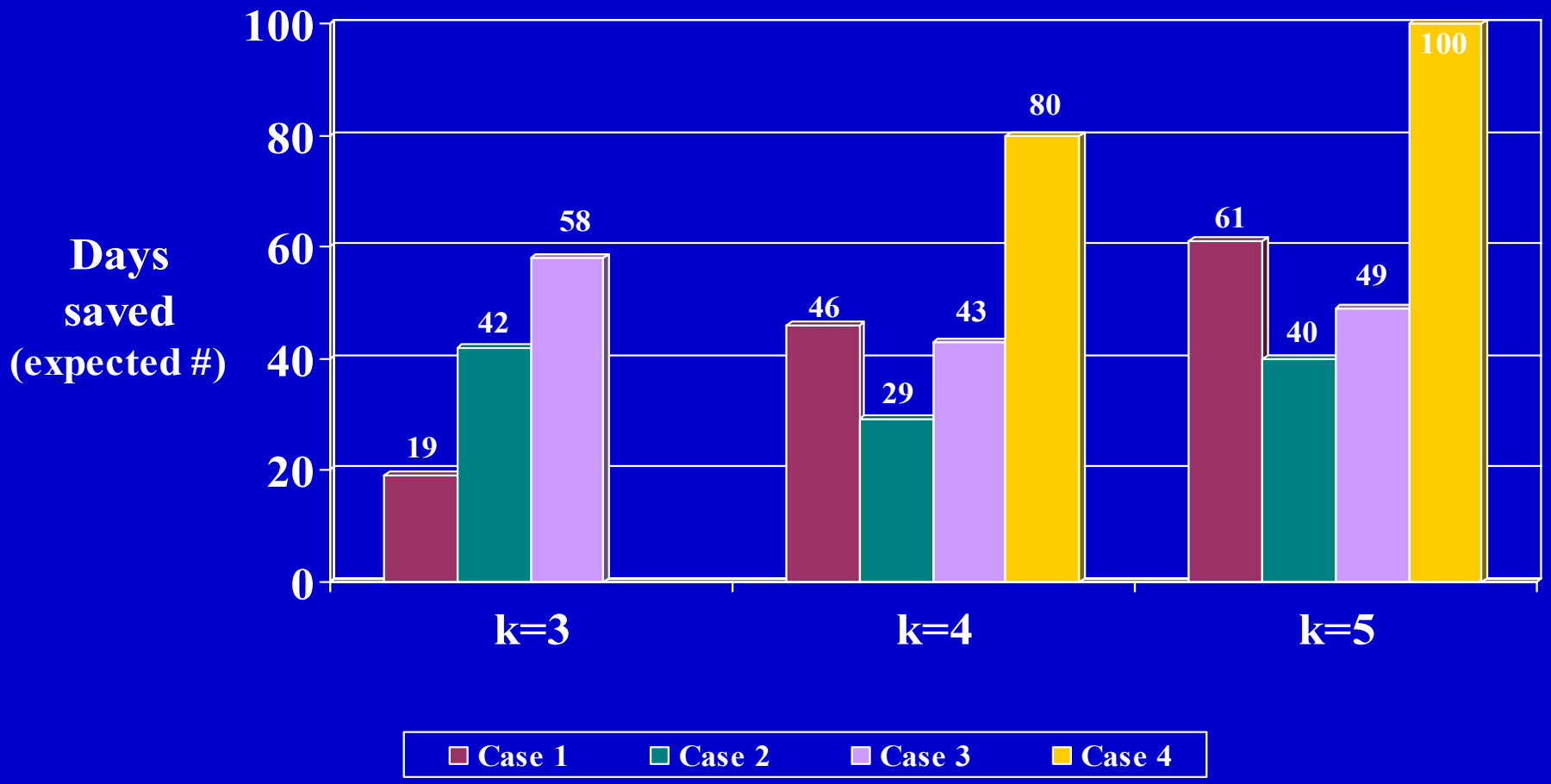

Note: We do not obtain a reduction in # of pts vs Scenario 2 in all cases with 60d, or 90d primary endpoint.

# Shortening of Timeline vs Scenario 1 (Phase II + Phase III always)




Note: Ph II/III also gives large time savings vs Scenario 1 in all 11 cases with 60d, or 90d primary endpoint.

# Shortening of Timeline through Phase II/III vs Scenario 2 (Phase II + conditional Phase III)




Note: Ph II/III also gives a time saving vs Scenario 2 in all 11 cases with 60d, or 90d primary endpoint.

# Further Work

- Extensions to decision rules when we keep 1 dose after interim
  - Allow for shape of dose-response curve in decision rule
  - Take account of safety problem chance at each dose
  - Allow for dose selection on early visit of primary endpoint, or on a surrogate (see Todd, 2003)
    - particularly useful when patient duration ≥ 90d
- Devise rules based on allowing 2 doses (& PL) to be kept at interim

# Conclusions & Practical Considerations

- Methodology presented allows combination of dose-finding and confirmatory stage within one study without stopping
- Allows for dose selection based on Phase III primary endpoint
- Allows early stopping if all doses ineffective

# Conclusions & Practical Considerations

- Type 1 error is controlled exactly at 0.025
- Rules should be fully pre-specified in protocol
- Use independent group to operate rules
- Discuss with FDA in advance if plan to conduct a Phase II/III combination study

# Conclusions & Practical Considerations

- If we start with 3-5 doses of test drug, then Phase II/III design gives high power
- Where this approach is consistent with needs of the program, Phase II/III design can:
  - **Cut Costs by Reducing Number of Patients**
    - **by 32% vs Scenario 1; by 3.5% vs Scenario 2**
  - **Cut Drug Development Time**
    - **by 289d vs Scenario 1; by 46d vs Scenario 2**

[Above results are medians of 11 cases considered, based on 30d endpoint]