Considerations in Planning Multi-regional Clinical Trials

H.M. James Hung, Ph.D.

Division of Biometrics I, OB

OTS, CDER, FDA

Presented in PSI-DIA Journal Club, April 5, 2011

Robert O'Neill

OB/OTS/CDER, US FDA

Sue-Jane Wang

OB/OTS/CDER, US FDA

Disclaimer

The views expressed in this presentation are not necessarily of the US FDA

Multi-regional clinical trial (MRCT):

Simultaneous conduction of trial for multiple geographical regions under the same trial protocol

Essentially serves two interrelated purposes

- assessment of global treatment effect
- use the trial results to bridge from global to local or between regions

Sources of regional differences

- intrinsic factors (ICH E5)

  race, genetic factors, …

- extrinsic factors (ICH E5)

  background treatment, social factors, health care system, medical practices, …

- quality of trial conduct or data

Values of MRCT

- Can yield a global effect estimate with best precision
- Global estimate may be best for bridging if effect estimates are similar among regions
- Offer opportunity to study regional differences of real interest
- Stimulate collaborative clinical research among regions for worldwide public health

Values of MRCT

- Raise awareness of concept of ‘quality’, can enhance trial quality for all local regions
- Harness global harmonization in trial standard
- Nurture clinical trial leadership w/ global view
- Cost effectiveness, ethical standard, regulatory standard, data/trial quality assurance, …..

Challenges of MRCT

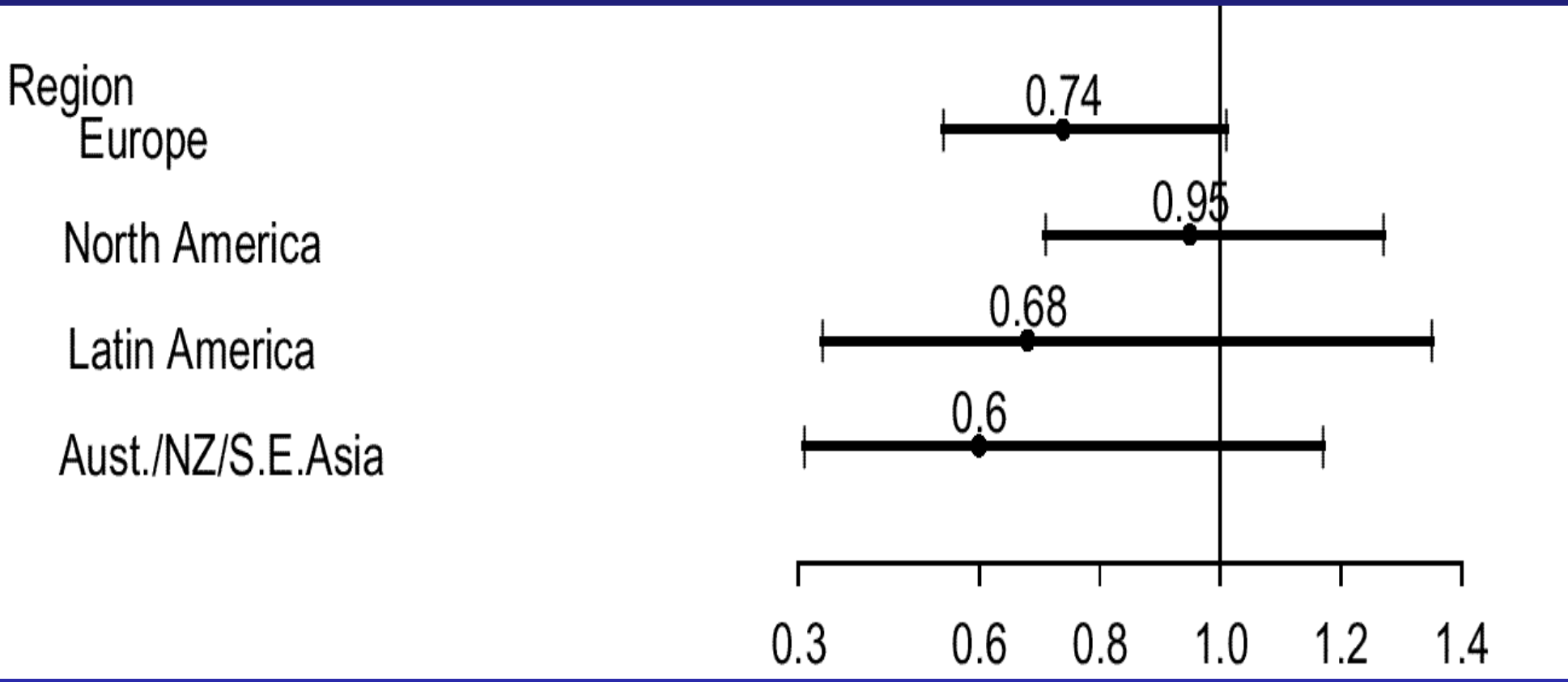
- Regional differences of effect estimates appear in many MRCTs
  - causes unknown
  - interpretation difficult
  - unclear about how to tease out real differences of interest from observed differences
  - unclear how to consider them in trial planning
  - how to best inform consumers is unknown

A number of NDAs showed possibly real regional differences in drug effects

- IDNT, RENAAL
- MERIT-HF
- Meta-analytic look of schizophrenia trials

…..

IDNT



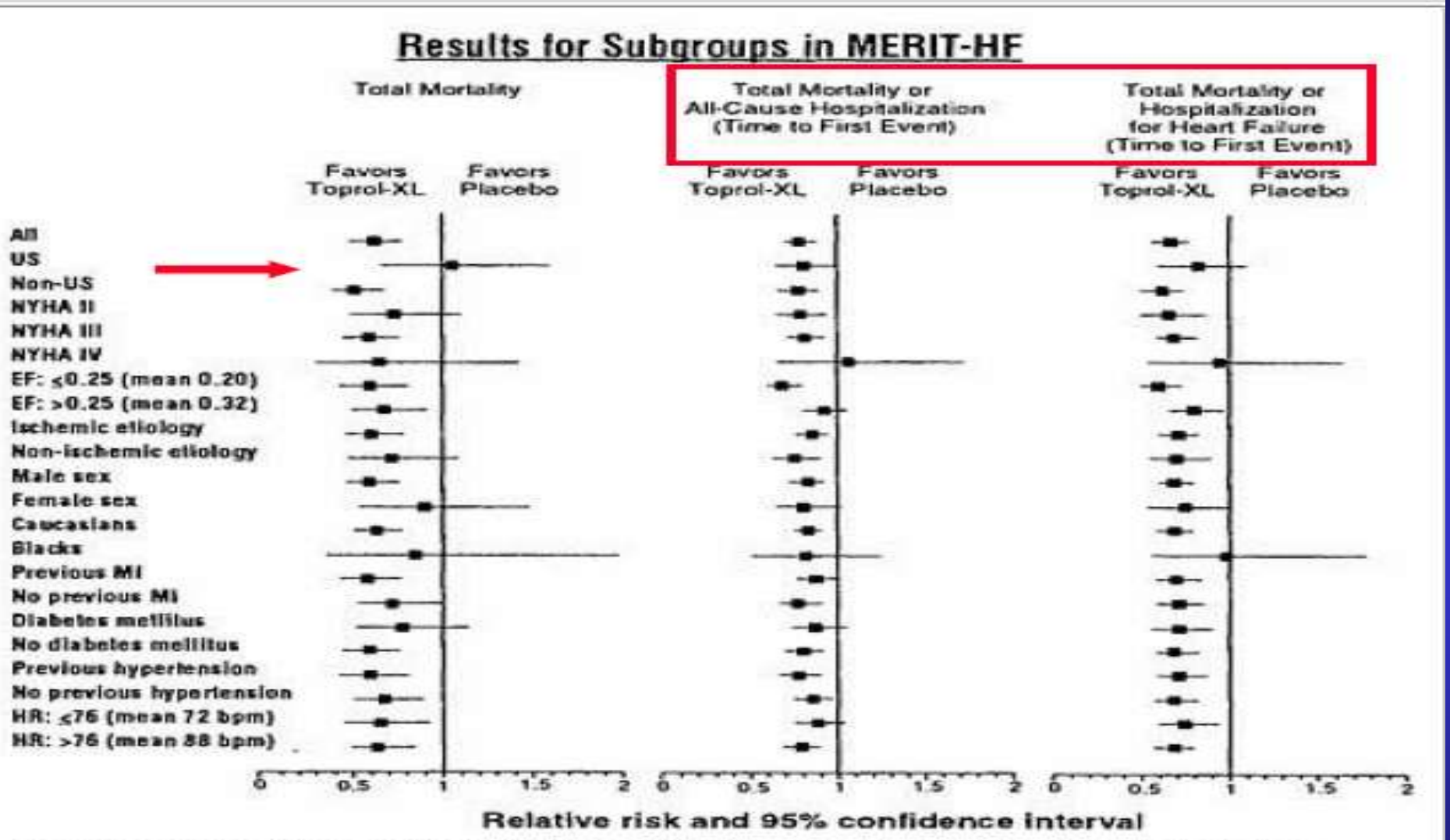
Relative risk (irbesartan/placebo) of DSC/ESD/D

RENAAL (interaction p = 0.044)

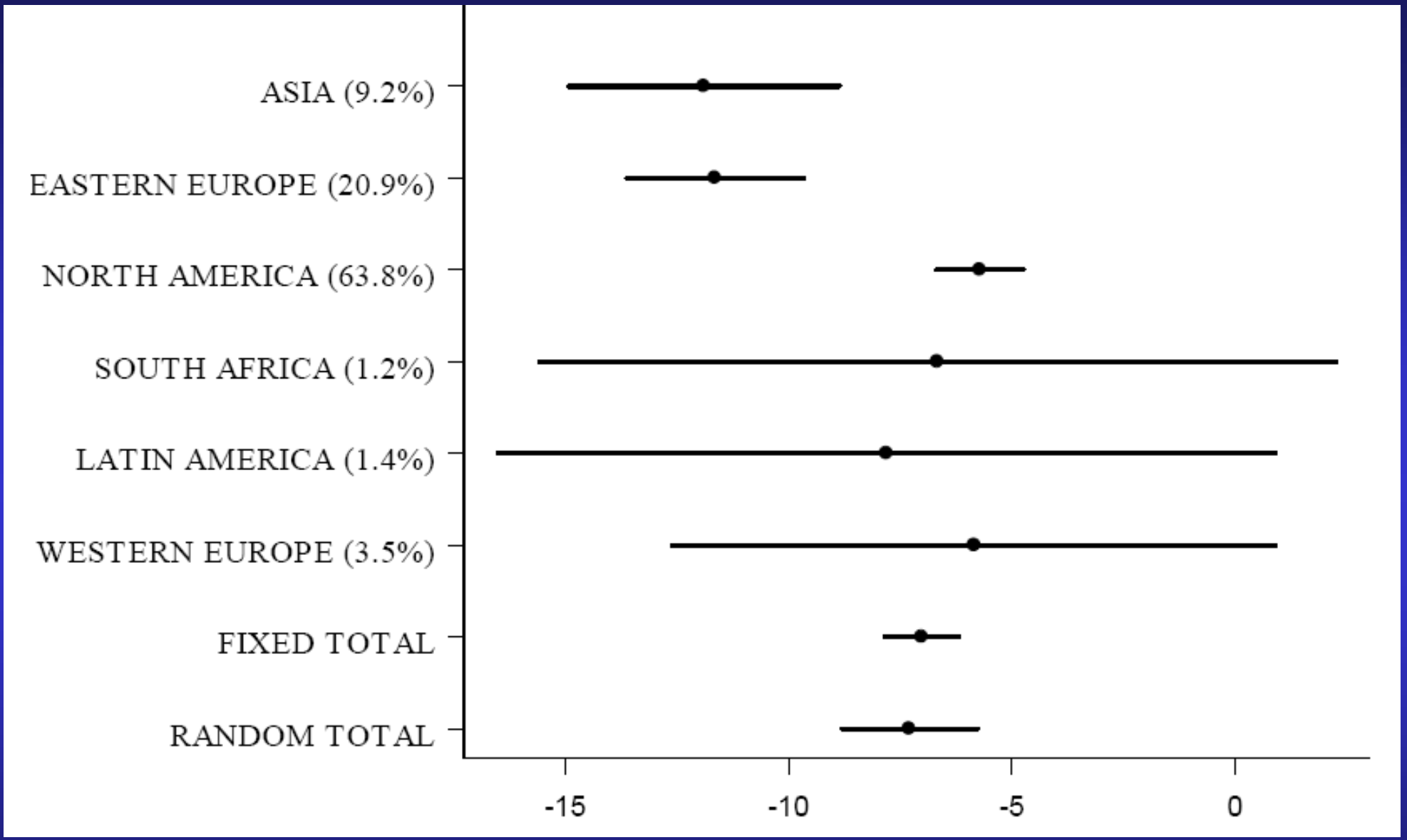
Region	TRT	Control	HR
Asia (17%)	39%	59%	0.55
Europe (19%)	38%	35%	1.05?,0.94?
Latin Amer(19%)	57%	58%	0.93
N. Amer (45%)	42%	43%	0.94
Overall	44%	47%	0.84 (p=.022)

HR: hazard ratio (losartan/placebo) of DSC/ESD/D

Qualitative or quantitative interaction?



O'Neill (2007)



Chen et al (2010, PST, about schizophrenia)

Challenges of MRCT

- Trial/data quality assurance

  - disparity in concept of ‘quality’

  - disparity in trial/data monitoring at local level

Challenges of MRCT

- Trial/data quality assurance

  - difficulty in trial/data inspection (translation, cultural aspects, resources, …)
  - regulatory enforcement authority, impartial, free from conflict of interest, sufficient resources, adequate role, ……

Considerations in Planning MRCT

- Endpoints culturally sensitive?

  If yes, multi-regional trial is not a good option

- Define ‘region’

  One definition is desirable

  Multiple definitions may be needed

  Consider defining it w/ intrinsic/extrinsic factors

- Implement quality measure in each region

  Concept of quality

- Consider consistency/inconsistency assessment in trial planning, e.g.,

  Japan MHLW (2007), Kawai et al (2008),

  Quan et al (2009), Uesaka (2009),

  Hung et al (2010), Ikeda & Bretz (2010),

  Marschner (2010)

- Explore possible need of more conservative sample size planning
  - Need prior experiences
  - Global estimate is still the key
  - Discuss extent of acceptable regional difference

Consistency Consideration - Design

Japan MHLW (2007): Meet the following "consistency" criterion

$$M1: \quad \hat{\delta}_1 \geq \pi \hat{\delta}_{all} \quad , \ \pi \geq 0.5$$

$$M2: \quad \hat{\delta}_i > 0 \quad , \quad \forall i = 1,\ldots,K$$

Have substantial implications on sample size distribution to the regions

Kawai et al (2008) consider *M2*

Minimum sample size for the smallest region such that

$$P(\hat{\delta}_i > 0,\ \forall i = 1,\ldots,K) \geq 1-\gamma, \qquad \gamma \leq 0.20$$

or

$$P(\ \hat{\delta}_i > 0,\ \forall i = 1,\ldots,K \mid \hat{\delta} > z_{\alpha/2} se(\hat{\delta})\ ) \geq 1-\gamma$$

For *K* = 3, the minimum sample size for the smallest region can be as low as 0.15N for γ = 0.20

Quan et al (2009) consider *M*1

$$\Pr\left(\ \hat{\delta}_1 \geq \pi \hat{\delta}_{all}\ \right) \geq 1-\gamma\ ,\quad 1-\gamma \geq 0.80$$

under $\delta_1 = u\delta_{all}$,

provided that total N is planned as usual

When $\pi = 0.5$, $u = 1$, $\alpha = 0.05$, $\beta = 0.1$, $\gamma = 0.2$,
$N1 = (22.4\%)\ N$

Q: If the criterion is employed by all *K* regions, will the total of N1 be N? No assurance

Hung et al (2010) consider evaluating

$$P(\ \hat{\delta}_i < 0 \text{ in } m \text{ of } K \text{ regions } \mid \Delta)$$

where Δ is the global effect, e.g.,

$\Delta = \delta$ (hypothesized global effect), 0.5δ, ...
$\Delta \equiv$ d (observed global effect)

If no regional difference in effect size Δ,

$$P(\hat{\delta}_h > 0 \mid \Delta) = \Phi(\Delta\sqrt{\lambda_h N})$$

Define $\pi_h = 1$ if $\hat{\delta}_h > 0$ and -1 if $\hat{\delta}_h \le 0$

$$P(\text{ m of K regions yielding } \le 0 \text{ effect } \mid \Delta)$$

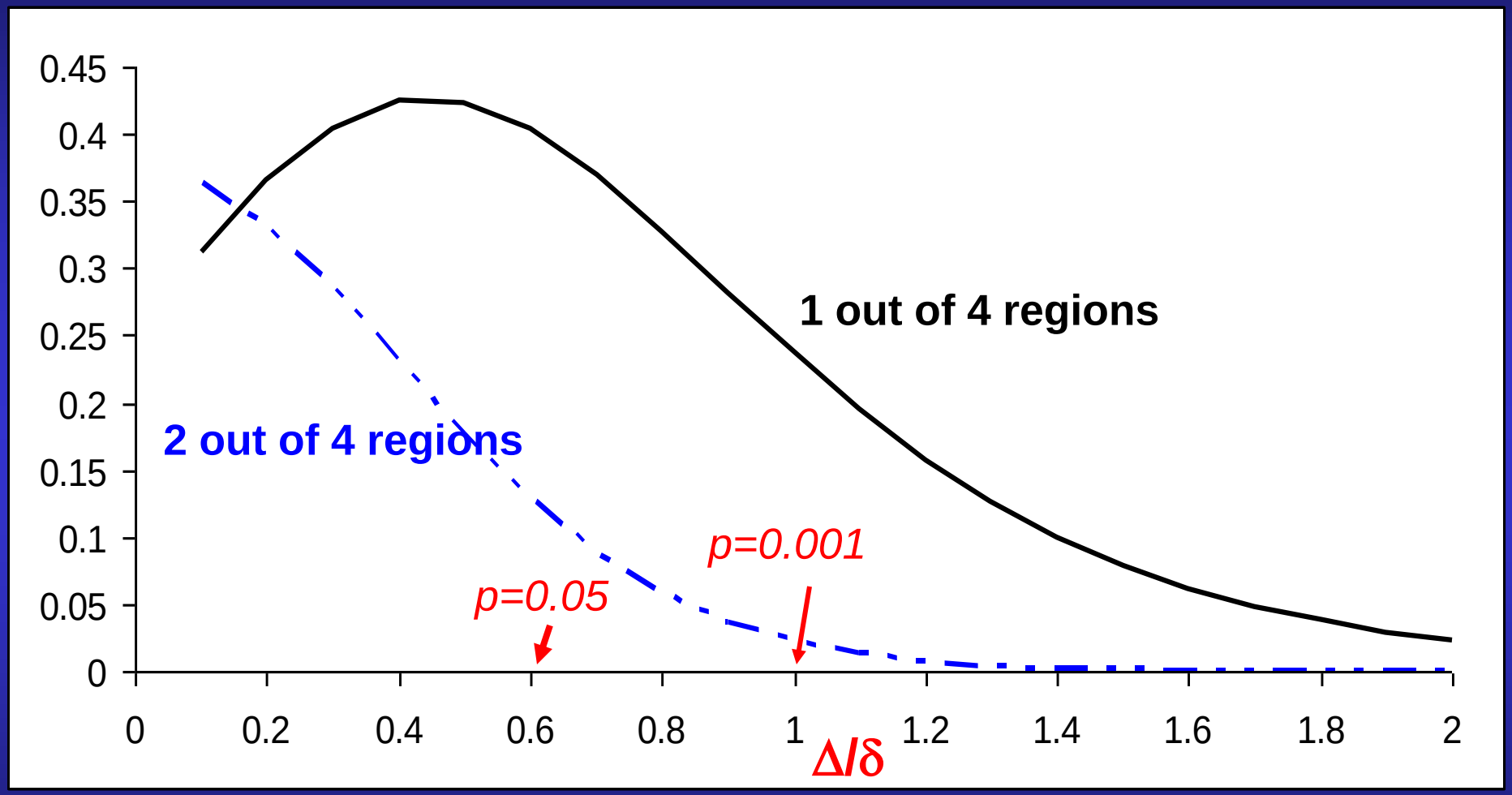
$$= \sum_{R_m} \prod_{h=1}^{K} \Phi(\pi_h \sqrt{\lambda_h}\sqrt{N}\Delta),$$

$$R_m = \{(\pi_1, \ldots, \pi_K) : \sum_{h=1}^{K} \pi_h = K - 2m\}$$

Example

Suppose a multi-regional (4 regions) clinical trial is planned to detect a postulated effect size $\delta > 0$ at 5% level of significance and power 90%, assuming all regions have an equal variance

P(m of 4 regions show nonpositive drug effect)



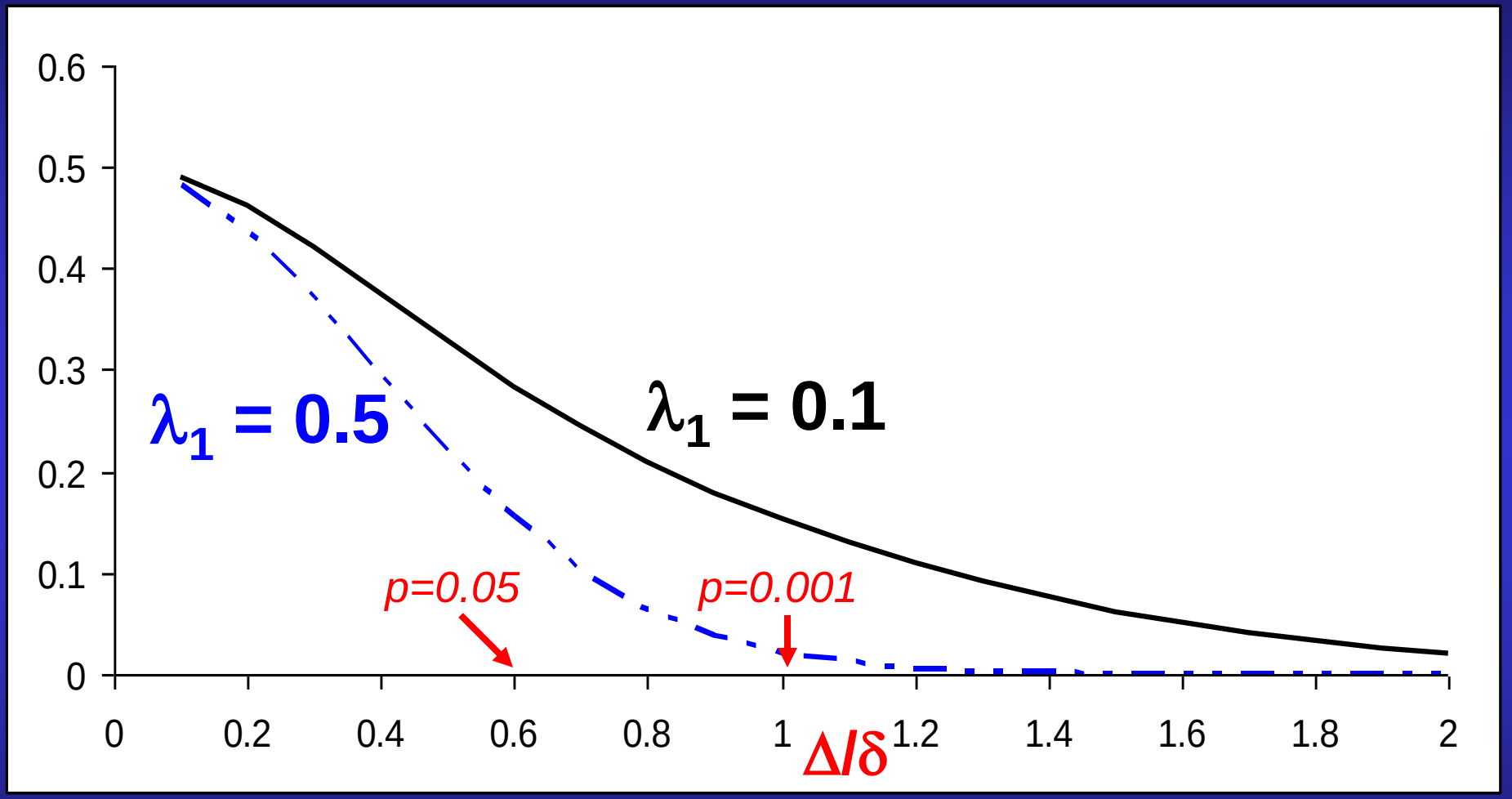
Sample size allocation to 4 regions = (.2, .1, .3, .4)

δ to be detected w/ 90% power

However, at the analysis stage, it is often to look at a specific local region (called Region 1) versus the rest

So it is also important to consider the probability of a local region showing a negative drug effect versus the rest for planning

P(Region 1 show a nonpositive drug effect)



Sample size allocation for region 1 vs. the rest =(λ_1, 1- λ_1)
δ to be detected w/ 90% power

Consideration in Sample Size Planning

Example: Five regions, drug vs. placebo

N distn: (20%, 10%, 40%, 10%, 20%)

To detect a global effect size $\Delta=\delta$ at 0.05 level of significance and 90% power

N: total sample size necessary

N_0: total sample size assuming $\sigma_\Delta = 0$ (consistent)

K geographical regions drug vs. placebo

n_h: total sample size of region h

$N = \sum n_h$ $r_h = n_h / N$

For simplicity, treat the problem as one-sample case

$$y_h \mid \Delta_h \sim \mathrm{N}(\Delta_h , \sigma^2/n_h)$$

$$\Delta_h \sim \mathrm{N}(\Delta , \sigma_\Delta^2)$$

Global estimate: $\hat{\Delta} = \sum_h n_h y_h / \sum_h n_h$

Should plan N to detect $\Delta = \delta > 0$ at level α & power 1-β,

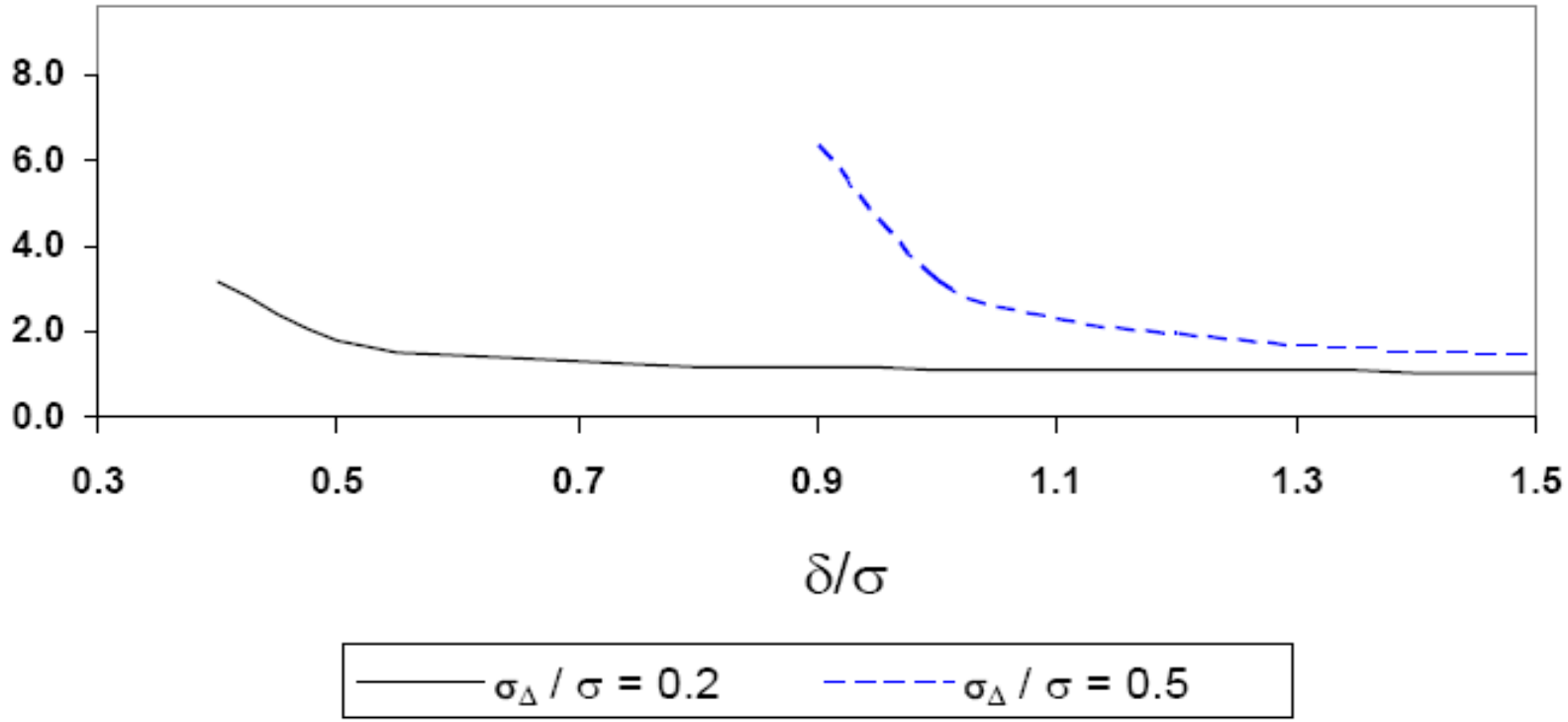
assuming $\sigma_\Delta \neq 0$

$$N = \left[\left(\frac{\delta}{\sigma(z_\alpha + z_\beta)} \right)^2 - \left(\frac{\sigma_\Delta}{\sigma} \right)^2 \sum r_h^2 \right]^{-1}$$

If, instead, $\sigma_\Delta = 0$ is assumed for planning sample size, then the resulting sample size N_0 may be too low. How low?

$$\frac{N_0}{N} = 1 - \left(\frac{\sigma_\Delta}{\delta}\right)^2 (z_\alpha + z_\beta)^2 \sum r_h^2$$

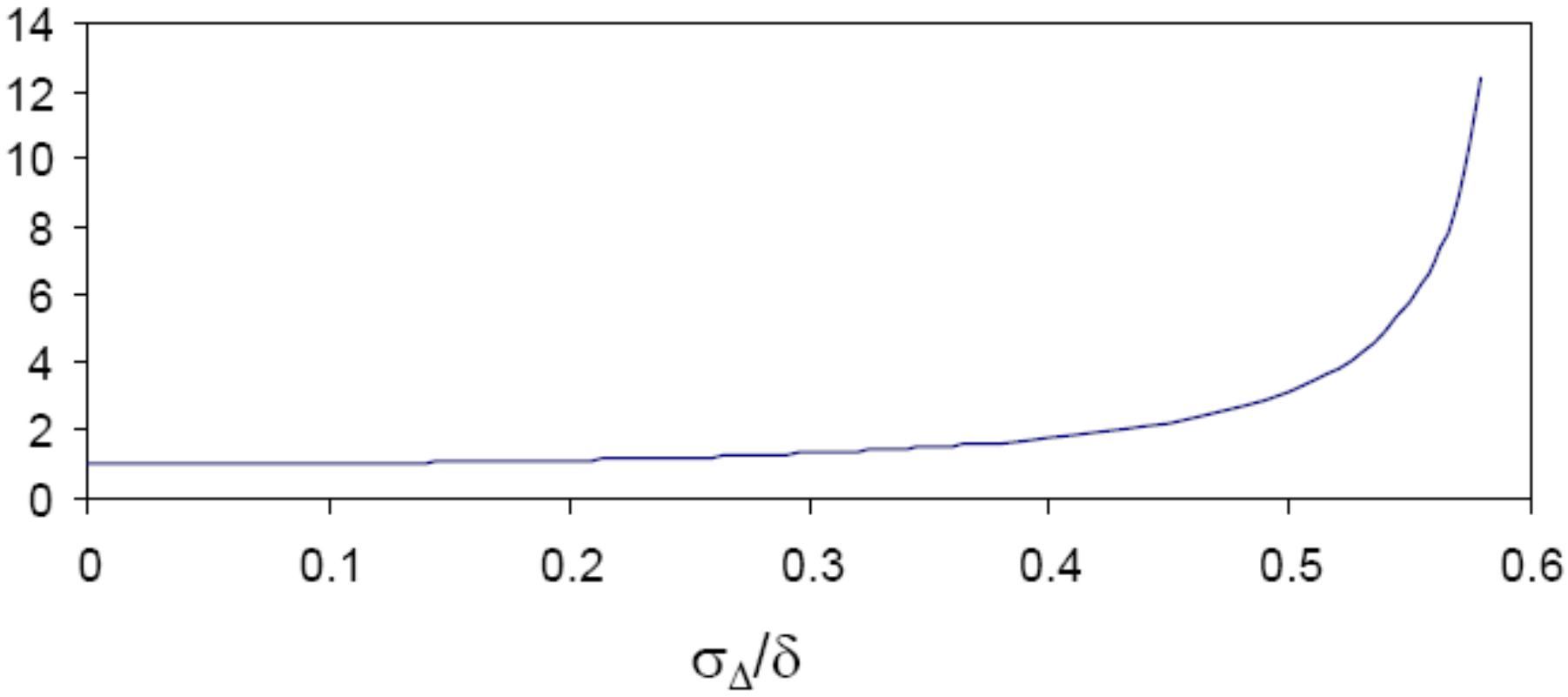
Figure 1. Sample size ratio N/N_0 versus (δ/σ)



α=0.025, β=0.1, K=5, $(r_1\ r_2\ r_3\ r_4\ r_5)$=(.2 .1 .4 .1 .2)

Hung et al (2010, PST)

Figure 2. Sample size ratio N/N_0 versus (σ_Δ/δ)



Hung et al (2010, PST)

Thank you for your attention!

Questions?