Plan and Analyze Multi-regional Clinical Trial: A Regulatory Perspective

H.M. James Hung, Ph.D.

Division of Biometrics I, OB

OTS, CDER, FDA

Presented in BBS Conference, Basel, Switzerland, September 16, 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

Two main goals:

- assess global treatment effect
- bridge from global to local or, if necessary, bridge between local regions

Why there are issues?

ICH E5: regional heterogeneity of treatment effect may be caused by:

- intrinsic factors (e.g., race, genetic factors)
- extrinsic factors (e.g., background treatment, social factors, health care system, medical practices)

Why there are issues? <cont'd>

Quality of trial conduct or data may also contribute to regional differences, but these differences are not a scientific goal and should be minimized

Heterogeneity caused by intrinsic/extrinsic factors is of scientific interest

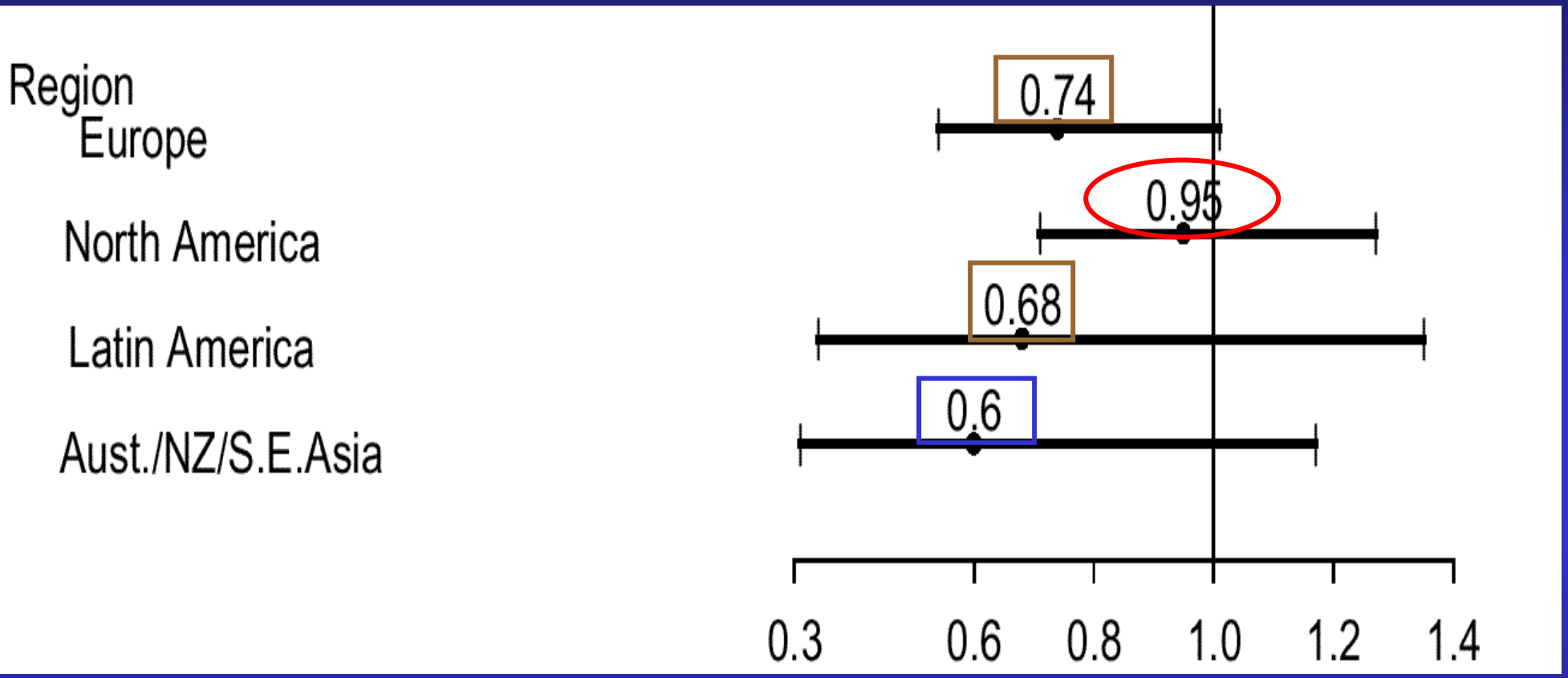
Why there are issues? <cont'd>

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

- IDNT, RENAAL (consistent heterogeneity)
- MERIT-HF (see no mortality effect in US)
- PLATO (see adverse effect in US)
- Integrated analysis of schizophrenia trials

…..

IDNT



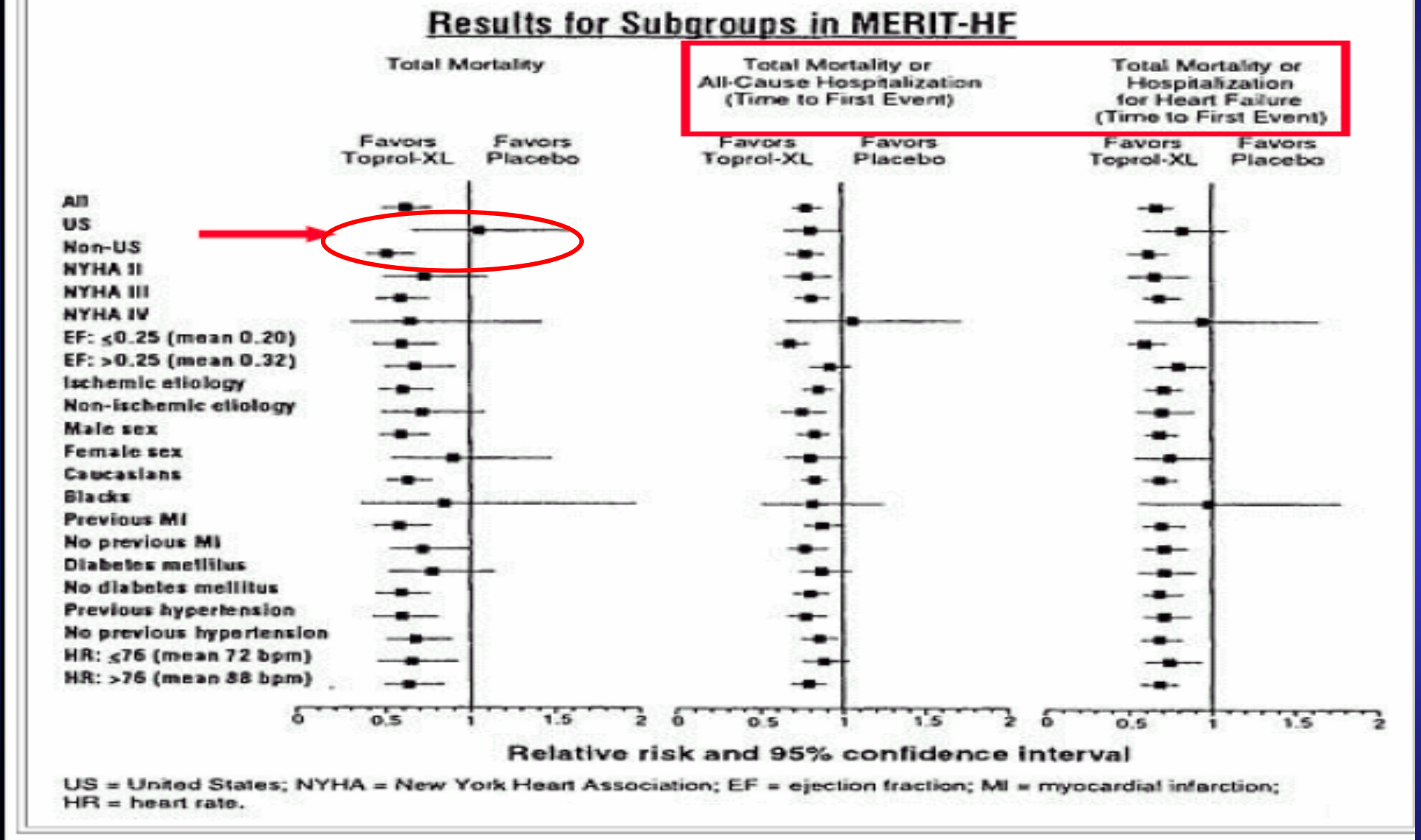
Relative risk (irbesartan/placebo) of DSC/ESD/D

RENAAL (interaction p = 0.044)

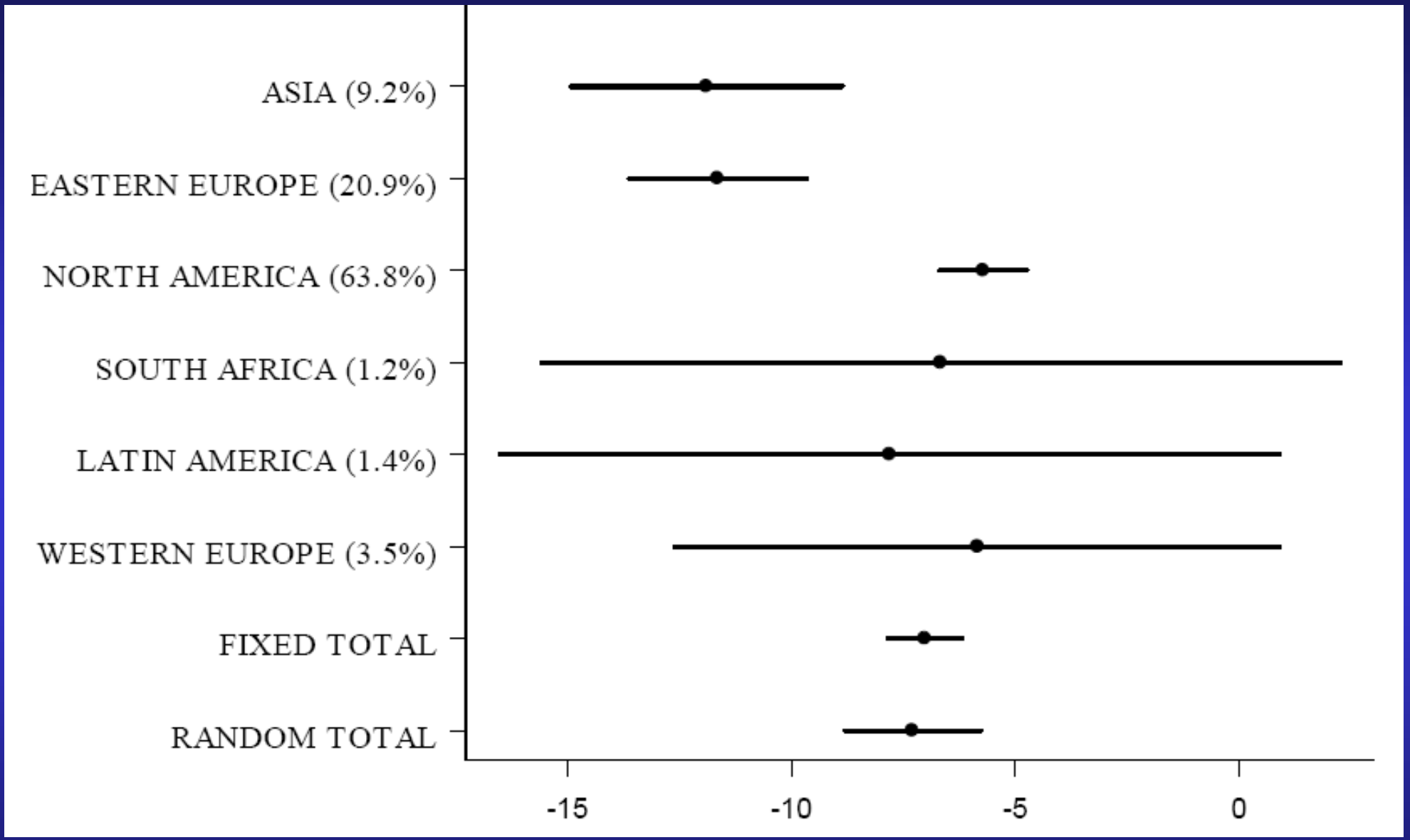
Region	TRT	Control	HR
Asia (17%)	39%	59%	0.55
Europe (19%)	38%	35%	1.05?,0.94?
Latin Amer(19%)	57%	58%	0.93
N. Amer (45%)	42%	43%	0.94
Overall	44%	47%	0.84 (p=.022)

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

Qualitative or quantitative interaction?



O'Neill (2007), drug label



Chen et al (2010, PS, about schizophrenia)

When regional differences of effect estimate appear,

- causes unknown
- interpretation difficult
- unable to extract real differences of interest from observed differences
- unclear how to consider them in trial planning
- how to best inform consumers is unknown

Considerations in Planning MRCT

- Endpoints culturally sensitive?

  If yes, MRCT is probably not a good option

- Define ‘region’

  One definition is desirable, but multiple definitions may be needed - sensitivity analyses

  Consider defining region using intrinsic/extrinsic factors

- Planning for implementing quality measure and monitoring trial conduct and data quality in each region

  Minimize regional differences in concept of quality, trial conduct and data qualities

- Consider consistency assessment and its impact on sample size planning and sample size allocation, e.g.,

  Japan MHLW (2007), Kawai et al (2008), Quan et al (2009), Uesaka (2009), Hung et al (2010), Ikeda & Bretz (2010), Marschner (2010)

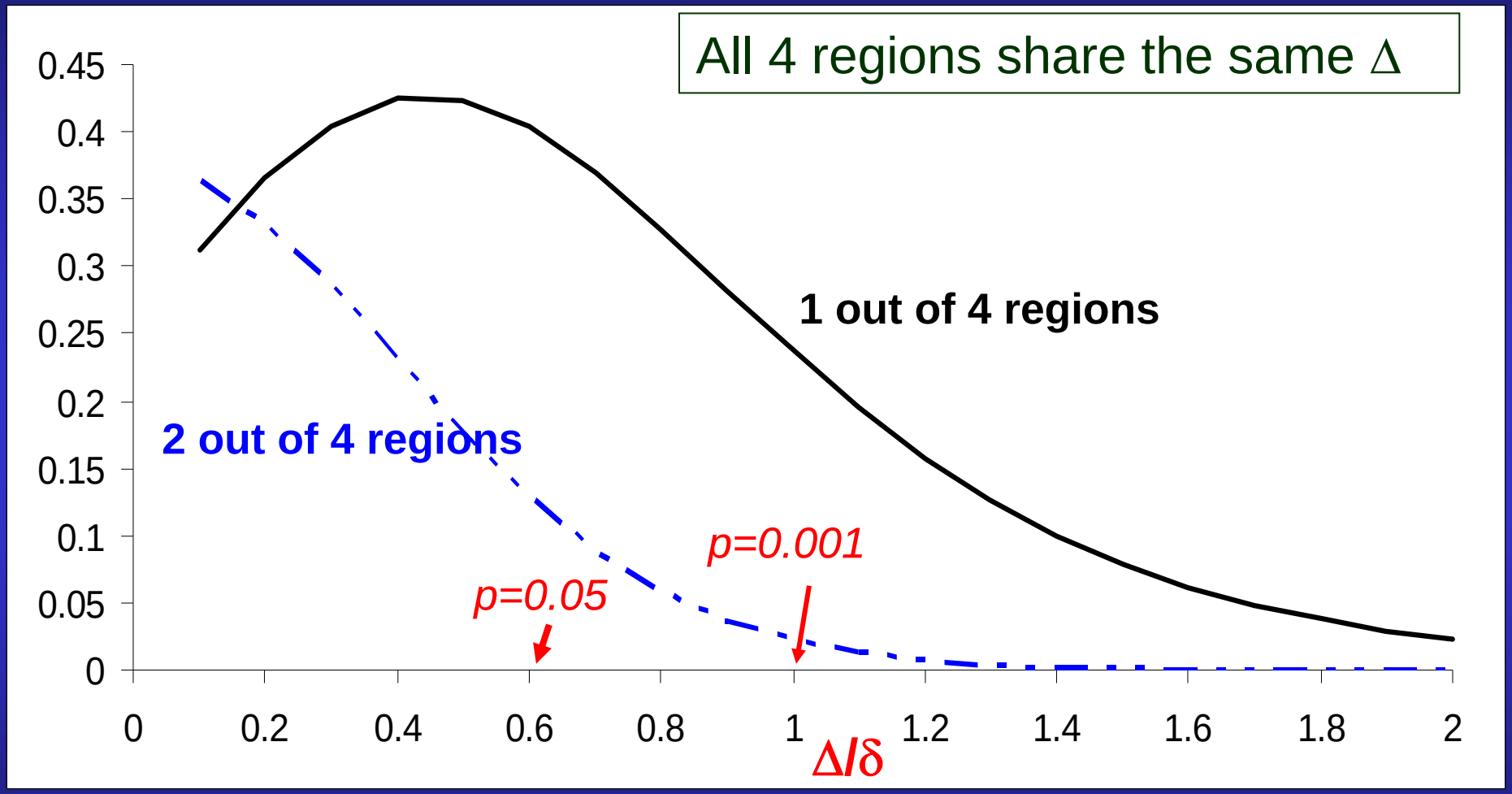
Implementing consistency measure may have large impact on sample size planning and allocation to regions

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 (w.l.o.g., assume standard deviation $\sigma = 1$)

Δ = true treatment effect shared by all 4 regions

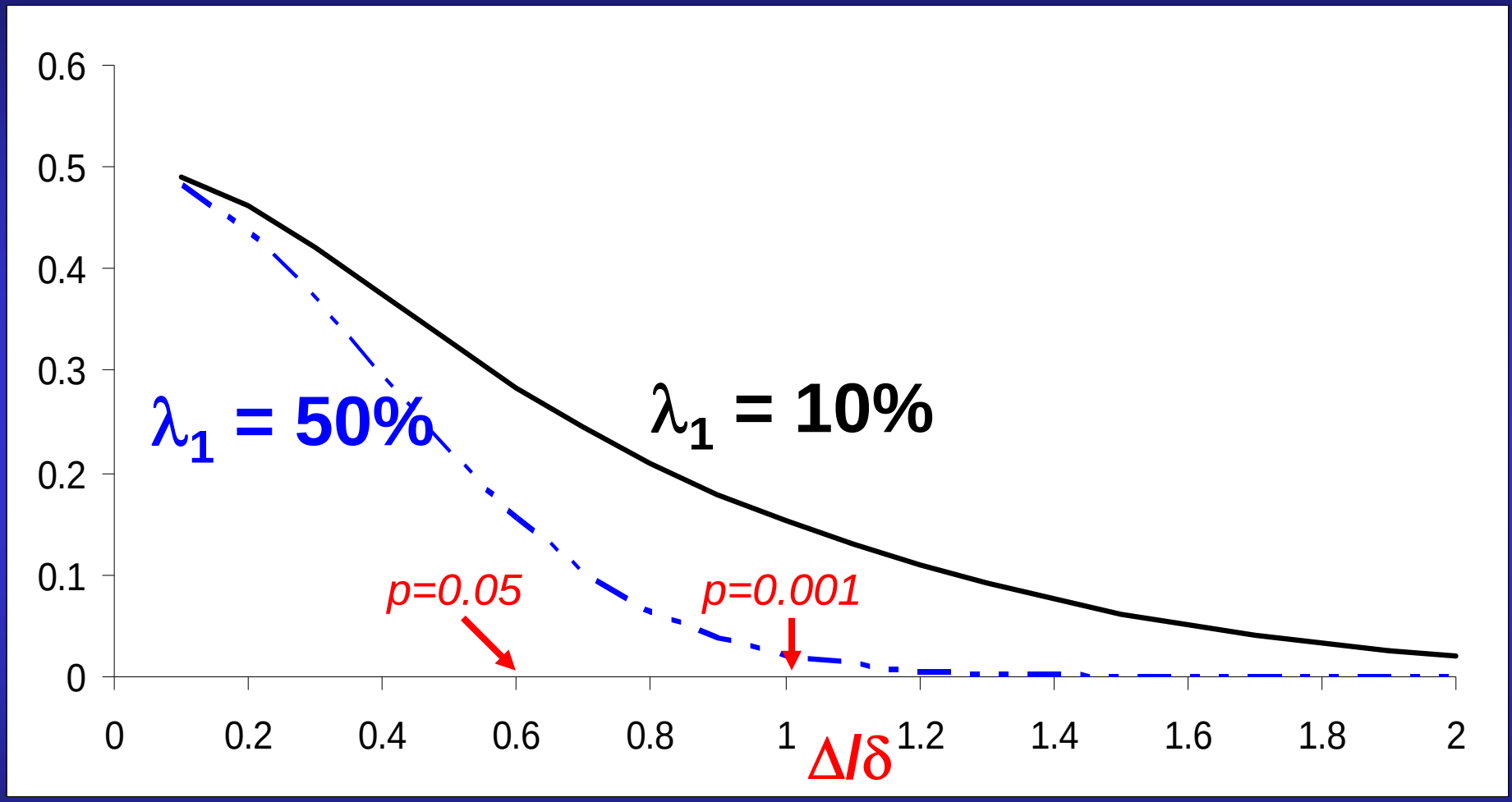
P(m of 4 regions show nonpositive drug effect)



Sample size allocation to 4 regions = (20%, 10%, 30%, 40%)

δ to be detected w/ 90% power

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

Probability of effect reversal may increase as number of regions increases or sample size allocation is more unbalanced

The smaller the sample size fraction for a region, the larger the probability of showing an effect reversal in this region will be

More sensible to strive for equal sample size allocation

- Explore possible need of more conservative sample size planning

  Need prior experiences

  Global estimate is still the key

  Discuss extent of acceptable regional difference

*Hung et al (2010, PS)

- Try to learn from external data base to identify covariates/factors that may cause regional differences and plan for obtaining the data of the covariates, e.g.,

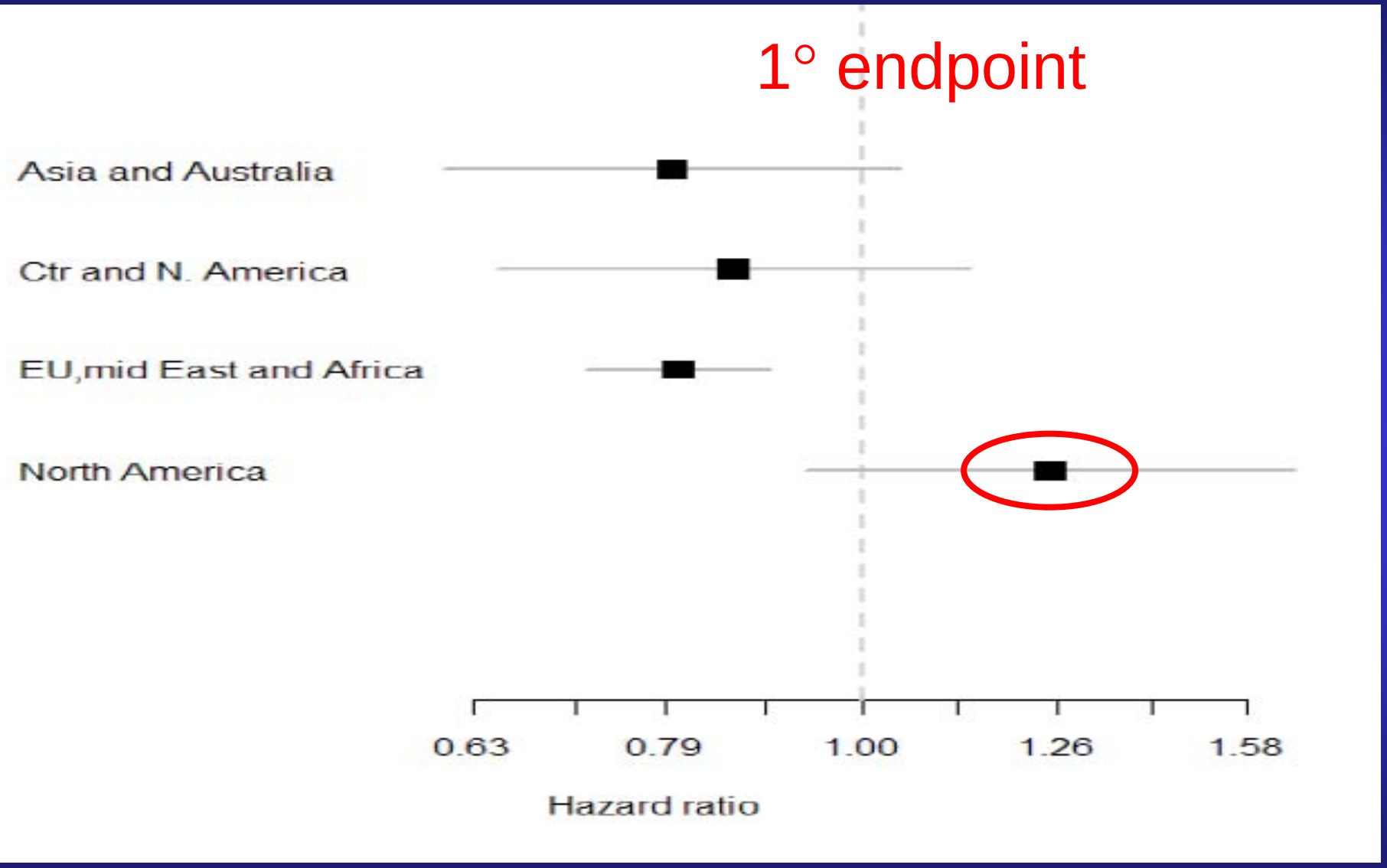
  Effect of intrinsic factor (e.g., ethnicity)

  Effect of extrinsic factor (e.g., medical practice, standard-of-care medications)

Considerations in Analyzing MRCT

- Global estimate is still the primary estimate representing each local region, unless large heterogeneity in treatment effects appears
- Explore possible sources for heterogeneity
  - intrinsic factors ?
  - extrinsic factors ?
  - data quality ?
  - due to chance ?

PLATO



FDA stat review by J.Zhang

No baseline covariates can explain heterogeneity

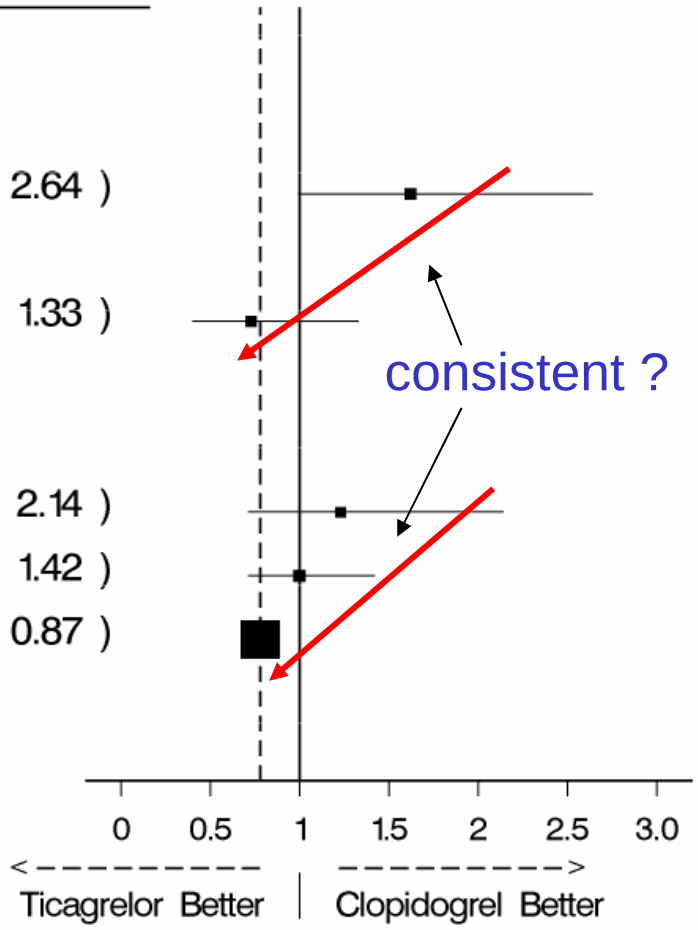
Concurrent use of aspirin appeared strongly associated with the heterogeneity, tons of analyses were performed to explore

Challenges:

- Concurrent use of aspirin not baseline covariate
- Aspirin use differs substantially between US and non-US
- US has < 10% of total events

Concurrent Use of ASA?

Region	ASA Dose (mg)	Ticagrelor N	Ticagrelor E	Clopidogrel N	Clopidogrel E	HR (95% CI)
US						
	>=300	324	40	352	27	1.62 (0.99, 2.64)
	>100 – <300	22	2	16	2	
	<=100	284	19	263	24	0.73 (0.40, 1.33)
Non–US						
	>=300	140	28	140	23	1.23 (0.71, 2.14)
	>100 – <300	503	62	511	63	1.00 (0.71, 1.42)
	<=100	7449	546	7443	699	0.78 (0.69, 0.87)



FDA AC presentation by J.Zhang

- If heterogeneity leads to searching for an estimate in lieu of global estimate for a local region, bridging needs to take into account covariates among the regions
- Shrinking regional estimates may need to be considered to avoid over- or under- stating treatment effect in local regions

  - need research

Remarks

- In considering use of MRCT for studying drug effect, need to determine whether global estimate can represent almost all regions
- Recognize an effect reversal from one of the regions may appear entirely or partially due to chance

- Once committed to MRCT, recognize global estimate is better than local estimate
- In reality, when heterogeneity appears, trial findings are difficult to convey to and can mislead consumers

- In planning MRCT, try to minimize the chance of apparent heterogeneity, e.g.,

  - more conservative planning (e.g., size trial for p-value of global effect to be at the level of 0.001 or less)

  - minimize imbalance of sample size allocation

- In planning MRCT, need to pay attention to substantial differences in trial conduct or data quality among regions

  Need to develop trial monitoring process that may differ in regions, taking into account cultural differences and others

- Plan for analyses looking for explanation for regional differences that appear

This work requires MRCT design to get complete data of covariates that may be effect modifiers

Back-up slides

Consistency Consideration - Design

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

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

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

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 $\pi = 0.5$, $u = 1$, $\alpha = 0.05$, $\beta = 0.1$, $\gamma = 0.2$,
*N*1 = (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, ...
$\Delta =$ 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 $\hat{\delta}_h \le 0$

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

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

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

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

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

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

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

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

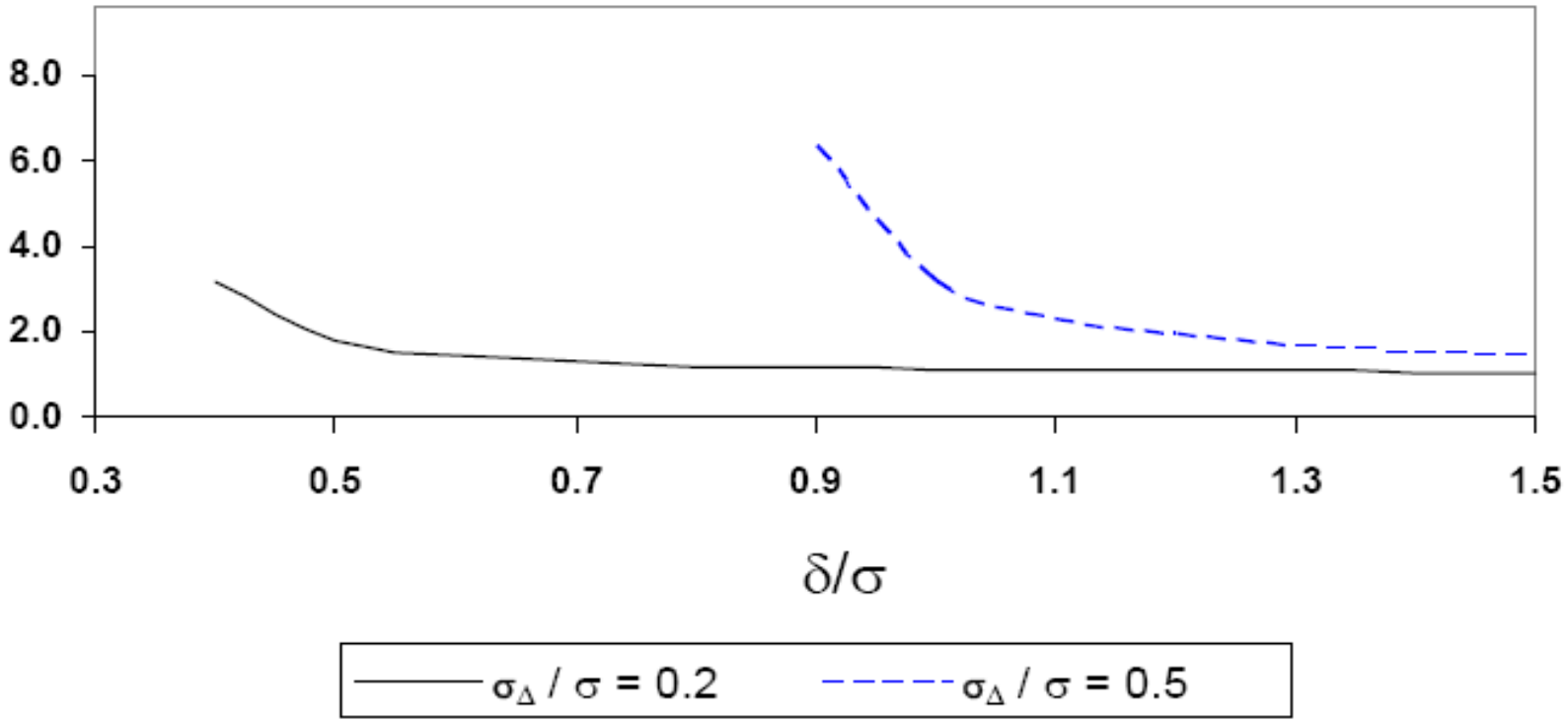
assuming $\sigma_\Delta \neq 0$

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

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

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

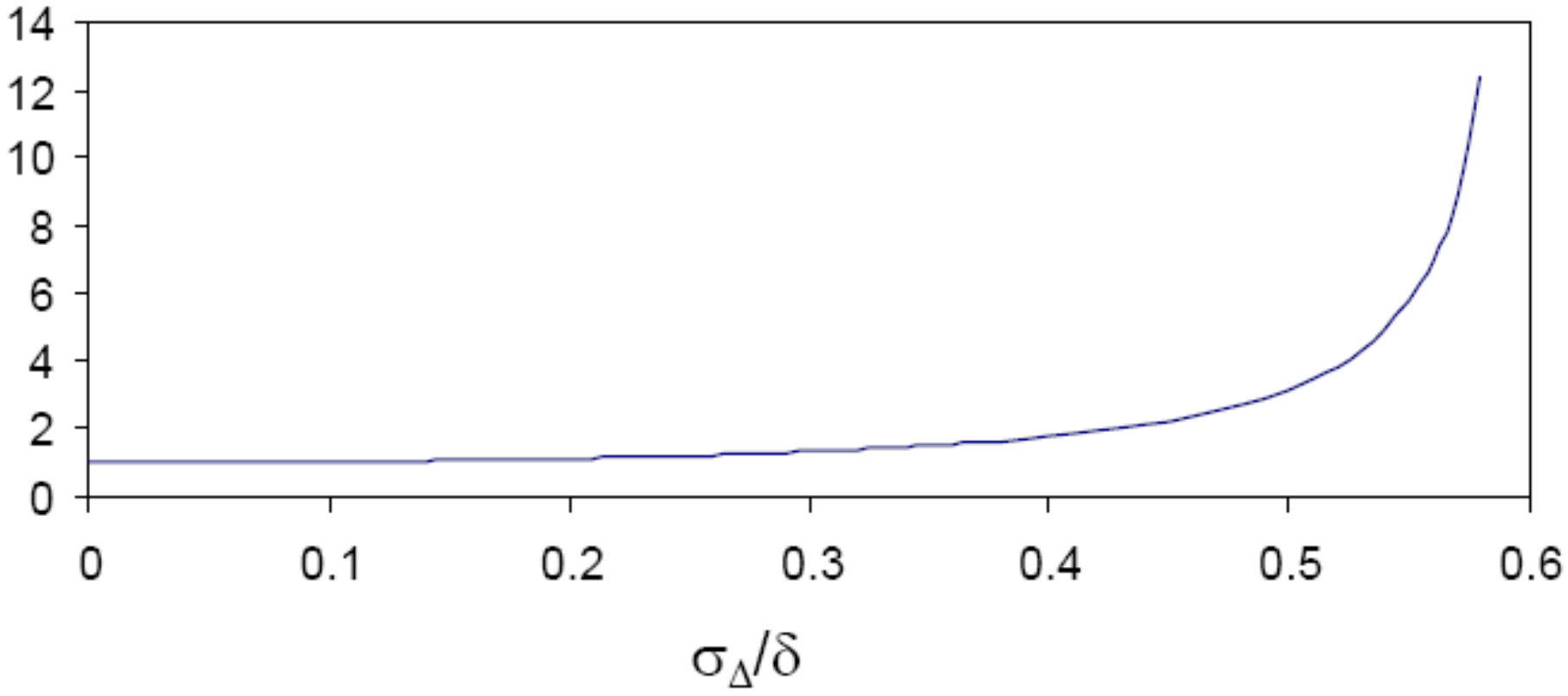
Figure 1. Sample size ratio N/N_0 versus (δ/σ)



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)